C Statement of Need

C.1 Background and Purpose

The mission of the National Oceanic and Atmospheric Administration (NOAA) is to understand and predict changes in the Earth's environment and conserve and manage coastal and marine resources to meet our Nations' economic, social, and environmental needs. NOAA's mission is embodied in its four strategic goals:

- Protect, restore, and manage the use of coastal and ocean resources through ecosystem-based management.
- Understand climate variability and change to enhance society's ability to plan and respond.
- Serve society's needs for weather and water information.
- Support the Nation's commerce with information for safe, efficient, and environmentally sound transportation.

In support of NOAA's mission and goals, NOAA conducts research and gathers data about the global oceans, atmosphere, space, and sun, and applies this knowledge to science and services that touch the lives of all Americans.

The ability to "predict changes in the Earth's environment" depends primarily on a diverse set of environmental models, usually requiring considerable computational resources. These models are developed through research and development (R&D) efforts within NOAA that occur primarily in three organizations: Office of Oceanic and Atmospheric Research's (OAR) Forecast Systems Laboratory (FSL), located in Boulder, Colorado; National Weather Service's (NWS) National Centers for Environmental Predictions (NCEP), located in Camp Springs, Maryland; and OAR's Geophysical Fluid Dynamics Laboratory (GFDL), located in Princeton, New Jersey. Modeling is becoming more prevalent throughout NOAA. Organizations such as the Pacific Marine Environmental Laboratory (PMEL), Climate Diagnostics Center (CDC), Aeronomy Laboratory (AL) and Air Resources Laboratory (ARL) are using models for ocean, climate, air quality and ecosystem modeling.

The purpose of this procurement is to provide the computational resources necessary to support continued advances in the environmental modeling capabilities and to meet other high-performance computing requirements that may arise within NOAA and at other partner agencies. This procurement will be known as the NOAA Research and Development High Performance Computer System (R&D HPCS) acquisition.

In order to fulfill the objective of awarding the R&D HPCS contract by late FY2005, NOAA will use the Department of Commerce's acquisition process referred to as "Concept of Operations" or CONOPS, described in "Department of Commerce - Acquisition Process Case for Change" (available from the CONOPS home page located at http://oamweb.osec.doc.gov/conops). The intent of this process is to create an acquisition environment that will benefit the Contractor and the Government. In order to successfully implement this process within this acquisition, the Government seeks the cooperation of the vendor community in an effort to conduct business in an atmosphere of integrity, openness, and fairness. It is essential that the Government acquires the best HPCS available for the budgeted level of funding and do so in an expedient and fair manner, therefore An R&D HPCS acquisition team ("the Project Team") has been formed within the Department of Commerce. A Project Agreement was drafted between the team and management to spell out the objectives, milestones, approach, budget and resources available for the project. The Project Agreement and associated documents describing this project are available on the Internet at the Project Web Site, http://rdhpcs.noaa.gov/. Interested parties should monitor this web site for additional information over the entire lifecycle of this procurement project.

C.2 Procurement Objective

The primary objective of this procurement is to acquire balanced, comprehensive computing capabilities to advance NOAA's research and development activities in environmental modeling. This includes not only high-performance computing but also additional hardware (storage devices and interconnects, communications interfaces, and other peripherals), software (primarily, but not limited to, the operating system, file system, storage management system, system utilities, performance and utilization software, diagnostic tools, and compilers), networking, full system maintenance, support services, and needed supporting infrastructure enhancements. This procurement will also support data management and archiving as well as analysis and visualization of model results.

This procurement embodies the new approach that NOAA has adopted to holistically manage its HPCS resources as a corporate asset. Prior to this procurement, NOAA organizations requiring HPCS procured them along organizational lines. NOAA determined that its HPCS resources will be established and managed based on functional requirements. NOAA established two functional requirements to meet its HPCS needs: (1) Operations and (2) Applied Research and Development. This Statement of Need describes the functional requirements of NOAA's Applied Research and Development.

NOAA's business goal is to achieve economies of scale in conducting its R&D HPCS acquisition while providing maximum flexibility in the resulting contract. NOAA intends to achieve this primarily through consolidation of requirements into fewer acquisitions. In order to achieve maximum flexibility, NOAA's intention for its R&D HPCS contract is to have a broad enough scope and to have sufficient reserve capability (contract ceiling) to accommodate both existing and unanticipated requirements.

Since computing is essential to NOAA's scientific objectives, the HPCS must be characterized by a very high level of reliability and availability. Availability of at least 96% (24 hours/day, 7 days/week, calculated each month) for the computational platforms has been the historical goal for NOAA's R&D HPCS.

The basic tenets and provisions of this Statement of Need establish what the Government feels are the minimum acceptable capabilities of the HPCS based on

NOAA's experience in performing its mission. However, innovation in proposed high-performance solutions is encouraged in addressing the needs of the Government. Newer technologies or an approach different from that presented here may provide opportunities to increase performance or enhance efficiency.

C.3 Current NOAA R&D HPCS

NOAA's current R&D High Performance Computing System is broken into complex sub-Systems that are designed to help accomplish NOAA's R&D missions. The NOAA R&D HPC sub-Systems are comprised of many components including: Large-Scale Computing, Development, Post-Processing and Analysis Computing, Storage, Archiving, Hierarchical Storage Management, Interconnects, Software, and Support. The current HPCS computational components include an IBM Power 4 cluster; a Linux cluster based on Intel Xeon processors interconnected by Myrinet; and SGI Origin 3800, Origin 3900 and Altix systems. The entire HPCS includes over 7.5 PB of storage capacity, >15 million files, and over a Gb/s in WAN connectivity. See section C.6.5.1 for a more complete description of the current HPCS.

C.4 Future Computing Environment

The resulting contract will allow NOAA to holistically manage its R&D HPCS resources as a corporate asset, striving to achieve, for each workstream, the most cost-effective balanced computing environment within available funding levels. The requirements for and desired features of NOAA's future R&D computing environment are expressed in this section independent of site location or resource configuration. The Government requires that the Contractor provide a configuration of HPCS components that maximizes performance while maintaining system balance.

The key to effective use of the HPCS is good system balance. Each component of the HPCS plays a critical role in maintaining the flow of information through NOAA's model simulations, analyses, visualizations, and ultimately into scientific insight and the dissemination of knowledge to the research community, to NOAA Operations, and to the other customers of NOAA's research. Consequently, the capabilities of the large-scale computers, hierarchical storage management system, analysis and visualization platforms, and network bandwidth shall be well-matched in a way that minimizes bottlenecks to the flow of information while maximizing performance. Achieving this proportionality in the acquired capabilities is an essential goal of this procurement.

Additionally, the computing resources available to NOAA must meet its scientific needs throughout the life of the contract, so the Government requires a phased delivery of all components of the HPCS. The initial delivery of the HPCS must provide a substantial increase over current capabilities in computational throughput for NOAA. At least one substantial upgrade to the sustained throughput must be provided during the base contract period, with archiving and other HPCS capabilities increasing commensurately. Although individual components of the HPCS need not be upgraded simultaneously, the Contractor is required to balance performance at all times.

The Government requires a single Contractor to be responsible for the design, installation, maintenance, and support of the HPCS that will advance NOAA R&D's modeling programs. The HPCS shall meet the stated objectives and specifications set forth in this SON and shall include all hardware and software necessary to operate as a complete, functional, balanced, and highly reliable system. A single Contractor will serve as the point-of-contact for the entire HPCS, even though the HPCS may involve components, from a number of different vendors, potentially located at multiple sites. Fundamentally, the Government must improve all aspects of NOAA's computing environment in order to fulfill its R&D mission. These aspects include:

- High-performance computing, including large scale computing, large scale data post-processing, and analysis capabilities. As described in Section C.4.7.5 and Section J of this RFP, the R&D workstreams are comprised of computationally intensive environmental modeling applications coupled to I/O-intensive codes and data storage.
- Hierarchical Storage Management System (HSMS). The HSMS shall provide archiving capacity to meet the expected rates of data production on the HPCS.
- Software for resource management, system administration, and application development. The Government requires operating systems and cluster software that can manage resources. Complete and functional FORTRAN90, C, and C++ application development environments shall also be provided.
- Reliability, availability, and support. Availability of at least 96% (24 hours/day, 7 days/week, calculated each month) has been the historical goal for NOAA's R&D HPCS. The HPCS must continue NOAA's historically high utilization of its computing resources. System reliability, availability, and Contractor support are considered fundamental aspects of the HPCS.
- High bandwidth connectivity to the NOAA operational HPCS.

This acquisition involves a Base Contract period of four years, and an Option Contract of an additional 4 years as described in C.6. The Government expects that initial delivery, for sub-Systems associated with workstreams 7-9, will occur in October 2005. The Government expects delivery, for any additional sub-Systems or components associated with workstreams 1-6, will occur in October 2006.

No less than 94% of the annual funding will be dedicated to the components of the HPCS specified in this RFP. Under direction from the Government, the remaining funds will be used to refine key areas of the HPCS or other aspects of NOAA's computing environment covered under the scope of the contract. Improvements will be made for performance, efficiency, or usability of the overall system. These areas may include, but will not be limited to, node, disk, memory, visualization, server, the network infrastructure, and additional support. The Government is not under obligation to direct this money to the Contractor. Key areas will be identified on an as-needed basis by performance assessments, including an annual system performance review by the Government. The Government and the HPCS Contractor will work together to identify the necessary items that will best meet NOAA's computing needs. Actual purchases for this purpose will be determined solely by the Government.

C.4.1 Large-Scale Computing (LSC) Component

The Contractor shall provide a Large-Scale Computing (LSC) component at a substantial increase in sustained throughput over NOAA's current supercomputers described in section C.7.1 below. Sustained throughput shall be measured by a throughput benchmark (section J.1) comprised of workstreams that are surrogates for NOAA's expected future workload. The scalability of the computational platform(s) shall be measured by a benchmark designed to reveal the performance and scaling characteristics of individual codes as they are executed on different processor counts.

The benchmark is comprised of a set of application workstreams. Each workstream is a sequence of steps designed to represent the complete end-to-end execution of a single modeling application used in NOAA R&D. Each workstream may contain multiple applications with a mix of compute-bound and I/O-bound codes. The Contractor shall minimize the execution time of each workstream and maximize the overall throughput when multiple workstreams are run concurrently, as described in Section J.

At least one substantial upgrade to the sustained throughput of the LSC, as measured by a throughput benchmark, shall be provided during the base contract period.

The Government requires resources for interactive work. The Government desires that interactive resources have a minimal impact on the batch production resources and vice-versa. It has been NOAA's experience that the nature of interactive R&D work creates resource contention with batch production jobs, and batch production jobs slow interactive response time. The Government desires the ability to reassign batch production resources for interactive work without a reboot or a restart to the batch system.

The Government desires the ability to test operating system and application software upgrades in isolation from the interactive and batch production resources on the HPCS.

Failover capability for job queuing and scheduling shall be provided. It is required that when any set of resources (such as disk or memory) in the LSC fails, batch jobs using those resources must be capable of being rerun without user intervention. In this situation, only interactive sessions hosted on the failed resources will be lost, and the sub-System must allow users to continue to be able to login interactively. It is desirable that failover be to processors that are binary-compatible with and running the same operating system level as the failed processors. The capability of the LSC to operate and be repaired in degraded mode is required. It is desirable that the LSC have no single point of failure. The Government requires an availability level (defined in C.9.2 below) of at least 96% on every major component in the R&D HPCS.

C.4.2 Development Component

For each LSC sub-System, a binary compatible interactive development platform shall be provided. The Development environment shall have the same software environment as the LSC sub-System.

C.4.3 Post-Processing and Analysis Component

Some of the HPC sub-Systems described in section C.7.2 below execute compute-bound and I/O-bound codes separately. For example, the climate computing resources, located in Princeton, use a Large Scale Cluster for the former and an Analysis Cluster, configured for very large sustained I/O, for the latter. The Contractor shall provide resources that can efficiently execute both types of codes, with the goal of minimizing the execution time of each workstream and maximizing the overall throughput when the workstreams are run concurrently.

C.4.4 Storage and Archiving Component

The Government's scientific data is a critical asset. As noted in section C.4.7, the Government requires the highest level of data integrity and at least 99% availability for access to its scientific data.

C.4.4.1 Home File System (HFS)

A Home File System (HFS) is required for each HPC sub-System provided. This file system will have small quotas and be used to store source-code, small data sets, and environment initialization files. The Home File System shall be globally visible to the specific sub-System to which it is associated. The Home File System shall be backed up utilizing an automated method (see section C.4.7.4).

C.4.4.2 Fast-Scratch File System

A Fast-Scratch file system is high-bandwidth local storage that is visible by all processors within a given batch job. The Fast-Scratch file system will be purged frequently. The Fast-Scratch file system will not be backed up. Initially, per sub-System, the Fast-Scratch file system shall be able to support 100 million files and individual files that are up to 100 GB in size. The Government expects that these requirements will grow over the life of the contract.

C.4.4.3 Long-term Scratch File System

Long-term disk storage is a staging area for users to temporarily place files that will continuously be manipulated. Large data sets that are not packed into a larger file are not conducive to a HSMS file system. This file system does not need to be backed up and can be purged of aged data. Initially, per sub-System, the Long-term Scratch file system shall be able to support 100 million files and individual files that are up to 100 GB in size. The Government expects that these requirements will grow over the life of the contract.

It is anticipated that the scientific data for WS4 – WS9 will be primarily accessible through the Long-term Scratch File System. The Government requires accessibility to its scientific data in the absence of the Large-Scale Computing resources.

C.4.4.4 Hierarchical Storage Management System (HSMS)

The HSMS component will be purchased. The Government requires a two-tiered storage scheme for its data archive, comprised of nearline storage (robotically mounted at high speed) and offline storage (robotically mounted with an emphasis on high reliability), that will satisfy the requests for scientific data that makes up NOAA's R&D workload (as discussed in section C.4.8). Multiple archives may be provided. If disk is required for the caching or staging of files, it shall be fault-tolerant and shall be provided in addition to the other file systems. For nearline and offline data, the Contractor shall provide a data recovery service to recover data in the event of media failure.

Initially, per sub-System the file system shall be able to support 100 million files and individual files that are up to 500 GB in size. The Government expects that these requirements will grow over the life of the contract. The Governments requires a peak positioning rate that appropriately balances HSMS performance with the rest of its associated sub-System.

$$P = \frac{(n*3600)}{(load + search + rewind + unload)}$$

P = peak Positioning rate
n = Number of tape drives
$load$ = time to Load a tape
$search$ = average Search time
$rewind$ = average Rewind time
$unload$ = time to Unload a tape

Each data archive shall be presented to the associated sub-System as a single /archive filesystem image. The Contractor shall provide all storage media used in the HSMS. The Government expects that files that haven't been accessed in one year will be migrated from the nearline storage to the offline storage. It is anticipated that the scientific data for WS1 – WS3 will be primarily accessible through the HSMS. The Government requires accessibility to its scientific data in the absence of the Large-Scale Computing resources.

Upgrades to the HSMS shall be provided commensurate with upgrades to the computational resources, with overall system balance the goal.

Within each archive, the HSMS software shall provide automatic migration between data archive tiers based on a combination of access time and file size. A method for determining the physical location of users' files within the storage hierarchy is desirable. User-specified migration between tiers through a single software interface is also desirable. Accounting on the HSMS that reports individual usage at the group and user level of all storage tiers is desirable.

C.4.5 System Integration

C.4.5.1 Transition to “One NOAA”

NOAA requires increased integration of its HPC resources over time in order to provide additional flexibility and robustness. Elements of the long-term vision include:

1) the ability of users to access any computational platform in order to:
    a. achieve code interoperability with good performance
    b. provide continuity of operations
    c. respond to changing programmatic requirements
2) the ability of users to access data from any platform of the HPCS and to migrate data from any existing archive to future archives
3) an ongoing collaboration with the Contractor to produce additional business processes that promote integration

Features of the NOAA R&D HPCS that promote this vision may include a single security profile, including single sign-on, and a single user interface into all HPC resources, which may be a single batch scheduler or a metascheduler. NOAA recognizes that not all elements of the vision may be available on the system provided during the term of the base contract, or may compromise performance unacceptably. In this case, NOAA may choose to forego implementation of one or more of the elements of this vision.

NOAA considers the risk of transition to the full suite of these capabilities at day one of the contract to be very high due to both the immaturity of the underlying technology for distributed operations and the significant business process changes required of users and institutions.

C.4.5.2 Wide Area Network (WAN) Component

NOAA has existing WAN capabilities as described in the Wide Area Networks (WANs) section of C.7. The vendor is responsible for providing any additional connectivity required by the location of a provided sub-System. For example, this requirement may be driven by data generation on the LSC, as described in section C.4.8.2, or between the post processing and analysis component and the data and archiving component.

NOAA’s Enterprise Network Target Architecture (ENTA) is currently being developed. The resulting NOAA-wide network infrastructure could be referred to as NOAAnet. Upon completion of this architecture, the Government may, at its option and at any time during this contract, require the Contractor to alter the WAN connection to comply with the NOAA ENTA. Such a change will be negotiated.

The Contractor is encouraged to be innovative in their Wide Area Network solution in order to optimize total system performance. The Government desires to balance the Wide Area Network solution that is provided with the impact it will have on the interactive experience of remotely located users.

C.4.5.3 High bandwidth connectivity to model and observation data

Near real-time observation and model data are required to be available for use on the NOAA R&D HPCS. The model data will be created and are available on the NOAA operational HPCS and much of the satellite data will also be available on this machine or the NESDIS CEMSCS computer. Note that the satellite data provided to the NOAA R&D HPCS must contain the same level of preprocessing as that provided to the NOAA operational HPCS. These data include future high-resolution satellite and radar data and future high-resolution regional and global model data. The following table shows the projected input data quantities required by each workstream throughout the contract period. Note that the data required by the workstream 4 (or 6) contain all data used in the other workstreams. The data required for workstreams 7, 8 and 9 are a subset of that required for workstreams 4 and 6. Thus, if multiple workstreams can access the observation and model data sets from workstream 4, only a single transfer of the information would be necessary. There may be additional needs for real-time data acquisition from other sites, but they would be funded externally to this contract.

Data Ingest (TB/day)

	WS1	WS2	WS3	WS4	WS5	WS6	WS7	WS8	WS9
FY2006							1	1	1
FY2007				3		3	2	2	2
FY2008				3		3	2	2	2
FY2009				6		6	4	4	4
FY2010				6		6	4	4	4
FY2011				12		12	8	8	8
FY2012				12		12	8	8	8
FY2013				24		24	12	12	12
FY2014				24		24	12	12	12

Approximately half of the data are model data. The other data are composed from operational and experimental satellite data which are expected to be used in data assimilation systems. The expected data sources are the current satellites still in use at the beginning of the contract possibly including METEOSAT, GOES-10-12 and NOAA-14-18 (both imagers and sounders) as well as NASA's EOS-Aqua and EOS-Terra satellites (MODIS on both and AIRS/AMSU-A on Aqua), SSM/I (imager data) and SSM/IS (F16) on the DMSP satellites. New polar orbiting satellite data are expected to come from NOAA-N', NPP (sounder and imager in 2007), SSMIS on F17 (2005), F18 (2007), F19 (2008), F20 (2010), METOP (sounders and imager in 2005), NPOESS (sounders and imagers) C1 (2010), C2 (2011) and C3(2013). For geostationary satellites, new sounder and imager satellite data will include GOES-N (2005), -O (2007), -P (2008) and – R (2012). All launch dates are approximate, and some need for transfer of simulated data streams of the same size as the actual data sets may occur up to a year prior to the launch dates.

Provision of the data to the NOAA R&D HPCS does not require the same timeliness as for operations. A delay of the data transfer of up to 6 hours (beyond that from the operations) is acceptable. Further delays of the data can be counted as machine downtime at the discretion of the government. Also, it is required that no significant gaps in the data sets occur because of communications and any delayed data should be communicated as soon as possible.

Currently available bandwidth should make the necessary data available within the Washington D.C. metro area and to the Skaggs building in Boulder. It is uncertain if the necessary bandwidth will be available to GFDL in Princeton. The costs for providing the necessary bandwidth for the communication of this data are part of this procurement.

C.4.6 Software Component

C.4.6.1 Resource Management Software

Efficient operation of the HPCS requires resource management that will facilitate use of the HPCS by NOAA's scientists as well as provide maximum throughput for their workload. The Government desires to implement a variety of charge-back algorithms for monthly processor time or enforce different resource allocations, including disk and tape quotas, for users, groups, or projects on each of the HPCS sub-Systems.

The Government requires resource management software that provides:

1. Batch queuing and scheduling
2. Operating system based accounting software, comparable to Unix SVR4 process accounting.
3. System activity monitoring software on the LSC that shows user and system CPU utilization, I/O wait time, and paging activity.
4. Disk quotas on disk space and number of files, comparable to Unix 4.3 BSD quotas.

The Government desires resource management software on the LSC that provides:

- Software partitioning on the LSC

C.4.6.2 Batch Queuing Software

The Government requires a Batch Queuing Software package to manage all HPC sub-System computing resources. The Batch Queuing Software shall be provided with documentation, training, and an established bug-fix process.

To maintain flexibility, the Batch Queuing Software shall support execution hosts with heterogeneous processor architectures and operating systems.

The Batch Queuing Software shall perform the following functions:

- Run parallel-processed jobs on processors dedicated to the job.
- Create a unique scratch filesystem temporary directory for each job.
- Provide a way for large input files to be pre-staged to the scratch filesystem before the job acquires dedicated processors, and for large output files to be post-staged after the dedicated processors are released.
- Enforce limits on per-process cpu time, job-total cpu time, and elapsed (wallclock) time.
- Track and display job-total cpu time as the job runs.
- Do priority-based job scheduling, allowing higher priority to be assigned to specific jobs, or all jobs owned by a user or group.
- Control the number of allowed CPUs, and maximum wall clock time by groups of users.
- Collect cluster-global job accounting information for each job, including:
  - the user and group
  - the number of processors dedicated to the job
  - the job-total cpu time
  - the elapsed (wallclock) time
  - the date and time of job completion

The Batch Queuing Software shall provide text-based displays as specified below. It shall be possible to create custom text-based displays, at a minimum by post-processing the display output with text-processing tools such as awk or perl.

- Display all cluster hosts, showing for each host
  - the host up/down status
  - the host load and resource usage
- Display all running jobs, showing for each job
  - the host(s) the job is running on
  - the number of processors dedicated to the job
  - the job-total cpu time
  - the elapsed (wallclock) time
- Display all queued jobs, showing for each job
  - the number of processors requested by the job
  - the date and time the job was submitted
  - the per-process cpu time limit for the job
  - the elapsed (wallclock) time limit for the job
- Display all job accounting records for a specified month

The following Batch Queuing Software features are desired:

- A single Batch Queuing Software package that provides batch job queuing, scheduling and execution for managing all HPCS computing resources.
- Uniform access to all HPCS computing resources, through use of a single Batch Queuing Software package.
- Meta-scheduling capabilities able to schedule batch jobs on multiple clusters over a wide-area network.
- A project or account feature with the ability to do priority job scheduling based on projects or accounts, and control user access to the high-priority projects or accounts.
- A job suspend or pre-emption feature that works for FORTRAN applications using the proposed parallel programming model.
- Priority scheduling of large jobs, keeping smaller, lower priority jobs from starting. In addition, during this draining process, scheduling of small, short jobs which will complete before the large job starts (backfill scheduling).
- Job scheduling based on consumable resources such as host memory, host disk space, or cluster-global shared disk space.
- System-level checkpointing.
- Interactive use integrated with the Batch Queuing Software that lets the batch execution environment run interactive batch job.
- Accounting of the cpu time used by interactive sessions.

The following Batch Queuing Software features are useful:

- Support for the user-level commands and command options of the POSIX 1003.2d Batch Environment standard (now part of POSIX 1003.1-2001).
- Integration with the Government's Linux desktop workstations.
- GUI-based displays.

C.4.6.3 COTS

The Government requires software licenses for the following COTS software. The Government desires that unrestricted licenses are provided.

- Matlab (http://www.mathworks.com)
- Mathematica (http://www.mathematica.com)
- IDL (http://www.rsinc.com/idl/index.cfm)
- S-Plus (http://www.mathsoft.com)

C.4.6.4 Programming Environment Software

Required commercial application software for the HPCS programming environment is comprised of:

1. FORTRAN 90/95, C, C++ programming environments, including:

a. ANSI standard FORTRAN 90/95, C, and C++ compilers
b. macro preprocessors
c. source-level debuggers
d. performance profilers
e. support for 64-bit IEEE reals and integers
f. support for reading and writing 32-and 64-bit IEEE floating-point formats in I/O operations
g. MPI-2 I/O, MPI-2 one-sided communications
h. MPI-1.1
i. The make utility

2. Facilities for source code management, including rcs and cvs.
3. OpenMP is required to the extent that processing units share memory.
4. Etnus TotalView parallel debugger.

Desired application software for the HPCS programming environments includes:

1. parallelized, optimized numerical libraries on the HPCS
2. optimized (possibly proprietary) I/O libraries.
3. data conversion facilities (for example, endian conversions and conversions of proprietary data formats used in the HPCS). Data conversion facilities will be required if the provided solution requires heterogeneous formats.

C.4.6.5 Community Supported Software

Required application software for the HPCS programming environments is comprised of:

- netCDF libraries and utilities (available at http://www.unidata.ucar.edu/)
- GNU make (http://www.gnu.org/software/make/make.html)
- Prior to the end of the Option Period, the Government expects that its modeling system software will become compliant with the Earth System Modeling Framework (ESMF). More information on ESMF is available at http://www.esmf.ucar.edu.

Required application software for interactive use is comprised of:

- Ferret (http://ferret.wrc.noaa.gov/Ferret)
- Ghostscript (http://www.cs.wisc.edu/~ghost/)
- Grace (http://plasma-gate.weizmann.ac.il/Grace)
- GrADS (http://grads.iges.org/grads/grads.html)
- ImageMagick (http://www.imagemagick.org)
- NCAR graphics (http://ngwww.ucar.edu)
- VIS5D (http://vis5d.sourceforge.net)

C.4.7 Reliability, availability, support

The HPCS shall continue NOAA's historically high utilization of its computing resources. Reliability, availability, and Contractor support are considered fundamental aspects of the HPCS.

C.4.7.1 Reliability

Downtime will be used to determine the actual System Life Throughput. Downtime is that period of time when all of an HPC sub-System's workload cannot be accomplished due to a malfunction in the Contractor-maintained equipment or software, or when the HPCS or one of its components is released to the Contractor for maintenance. Periods of Remedial and Preventive Maintenance count as downtime. Null time is that period of time when the workload cannot be accomplished due to environmental failure, such as loss of electric power or cooling, or recovery from environmental failure. . Null time will not be counted as downtime. Downtime for each HPCS component is based on the fraction of the resources available for that component's workload. It is arrived at through consultation between the Government and the Contractor, and determined ultimately by the Government. Downtime is accumulated on the HPCS if the Government is not able to execute its workload when access to the scientific data via the storage or archiving components is unavailable.

A component's downtime shall commence at the time the Government makes a bona fide attempt to contact the Contractor at the designated point of contact, (see section C.4.7.3). At this time the Government will begin a log of the problem which will be completed and signed by both the Government and the Contractor when the problem is resolved. Information to be entered into the log will be determined by the Government.

A component's downtime shall exclude any time in which the Government denies the Contractor maintenance personnel access to the malfunctioning hardware and/or software. Routine procedures for entry to a facility shall not be construed as denial of access.

A component's downtime shall end when the computer is returned to the Government in operable condition as determined by the Government, ready to perform all of the workload.

Preventive Maintenance (PM) is to be completed at times determined by the Government.

The testing and installation of every major operating system release installed at the request of the Government and one (1) minor operating system release installed at the request of the Government during any annual period will count as downtime. Preparation for and execution of post-upgrade LTDs, including any benchmark development by the Contractor, associated with the agreed-upon upgrades proposed at contract award, will count as downtime.

C.4.7.2 Availability

All components of the HPCS must perform as an integrated system to provide the Government with at least 96% availability. These components include not only those typically associated with a high performance computing system but also any WAN component provided as part of the HPCS. Additionally, the Government requires 99% availability for access to its scientific data. Scientific data archives may be on the Long-Term Scratch component or the HSMS component, depending on the overall architecture associated with a given workstream. The Government's data are critical assets associated with the HPCS. With the large amount of data associated with the HPCS, the Government requires the highest levels of data integrity.

Although availability will be monitored daily, it will be measured on a monthly basis. The Contractor shall establish and maintain accurate system monitoring and/or logs to support system availability measurements. Monitoring tools must be made available to the Government. Throughput shall be determined by relating system availability to the workstreams normally run on that system. For example, if sub-System A typically runs workstreams 1 – 3, workstreams 4 – 9 will not be used to determine the amount of "lost" throughput, if any.

Availability shall be determined by computing the ratio of total computation processor hours available for execution of R&D jobs to the total computation processor hours each month, with regards to Null Time. The time a computation processor is available for execution is determined by subtracting processor downtime from wall clock time. Spare processors can be included in the computation pool to reduce downtime, but downtime accumulates until the spare is made available in the system for job execution. Accumulated computational cycles (in CPU-hours) that are lost when jobs are lost due to component failure or component reboot will not count toward the system-life throughput calculation. If the accounting software cannot report the accumulated computational cycles for each active job at the time of failure, it will be assumed that four (4) CPU-hours were lost for all processor(s) on the failed system.

System Life Throughput, for a given workstream i, can be calculated by the following equation:

$$SLT_i = \sum_j \frac{T_{i,j} A_{i,j}}{B_{i,j}}$$

SLT = System Life Throughput
T = total wall-clock Time during system configuration j
A = proposed Availability
B = Workstream Benchmark time
i = WS number
j = system configuration period

An example of System Life Throughput for a sample workstream can be demonstrated for FY2001-FY2003. This calculation is based on actual calendar days beginning October 1, 2000. For example, if the proposed LSC throughput benchmark execution time is 10,800 seconds initially, and is upgraded to 7,200 seconds on October 1, 2001, and a 98% availability is proposed through the three-year period, the system life throughput SLT is given below (Note: there is 86,400 seconds in a day).

$$SLT_{sample} = \frac{((365\frac{days}{year})(86400\frac{s}{day}))(0.98)}{10{,}800\frac{s}{benchmark}} + \frac{(2(365\frac{days}{year})(86400\frac{s}{day}))(0.98)}{7200\frac{s}{benchmark}}$$

$$SLT_{sample} = 11446.4 benchmarks$$

Proposed throughput benchmark performance levels will be combined with the proposed availability level (98% is shown in the example above) to determine a measure of overall proposed System Life Throughput for the workstream. The actual throughput will be measured on a daily basis by monitoring the availability of the components associated with the given workstream. All performance levels must be met for each measurement of actual throughput regardless of past delivery of suites.

Although system throughput shall be monitored daily, the accumulated System Life Throughput of the workstream shall be calculated monthly. Shortfalls shall made up with new equipment brought in at no additional cost. Using the demonstrated benchmark performance on the upgraded HPCS, the Government will calculate how long the upgrade shall stay in place to compensate for the shortfall in throughput. It is the Governments goal to meet the toal System Life Throughput by the end of the contract in a manner that does not require frequent disruptions, front-loading of cycles, or back-loading of cycles.

At the option of the Government, shortfalls in throughput on the HPCS due to downtime shall cause downtime credits to accrue. These downtime credits shall be in lieu of bringing in new equipment. Downtime credits shall also accrue on all sub-System-wide available file systems.

To better reflect NOAA’s computational needs over time, changes in the HPCS workstream benchmarks shall be made by mutual agreement between the Government and the Contractor through the life of the HPCS.

The Government requires a credible method of maintaining system availability. The Government expects that fail-over and high reliability components will be used. The Government requires sufficient power during environmental failure to gracefully shut down all components of the HPCS. Further, the Government requires adequate power conditioning to insulate the System from power spikes and sags. It is expected that uninterruptible power supplies (UPSes) and power distribution units (PDUs) will be needed for all components of the HPCS to meet these requirements.

C.4.7.3 Support

The Contractor shall provide the Government with a designated point of contact to request maintenance. The Contractor shall maintain escalation procedures that allow the Government round-the-clock telephone contact with knowledgeable Contractor staff should the designated point of contact be unavailable.

For each provided HPC sub-System, the Government requires comprehensive support in order to meet the 96% availability requirement. Support professionals can consist of systems analysts, hardware engineers, and applications analysts. Contractor system analysts shall work closely with Government system administrators. Government system administrators shall have root access to the HPCS computing platforms. The Government requires a senior applications analyst to be available to the primary users of the workstreams for the duration fo the contract. The primary users are described in the User Profile, Section C.4.8.5. NOAA expects to provide offices for the senior analysts at its facilities at FSL, GFDL, and NCEP. It is anticipated that additional senior application analysts may be required during the transition to new technologies.

The Government requires an itemized list of all Contractor-supplied hardware and software items, and documentation of these items, in printable electronic form. Training shall be provided for NOAA computer specialists and operators in:

- system administration and tuning
- hardware operation and system overview

Training shall be provided for a large number of NOAA applications programmers in:

- application and shell programming
- programming languages and tools
- HSMS software user interface
- optimization

The Contractor shall provide the Government with a list of additional potential training topics. Training must be collocated with users to the greatest extent possible. On-line training with access to experts will be considered.

The Government desires to begin training when pre-delivery access to systems similar to those proposed for the HPCS is granted.

C.4.7.4 Automated backup

Automated backup shall be provided for all unique system images on the HPCS; each Home File System (HFS); and the inodes, metadata, and staged files of the HSMS. Software that allows users to restore /home files via a graphical interface is desirable.

For the HFS, a combination of full and incremental backups shall be done to robotically mounted tapes. These backups shall make it possible to restore files to their state on any day during the previous two calendar months. Minimal impact of these backups on the network load is desired.

A history of bimonthly full backups, for the previous 12 months, shall be produced for shelf storage. A history of annual backups shall be produced until the end of the HPCS system life for shelf storage. At the Government's discretion, shelf storage may be at a remote location. It shall be possible to restore files from these backups until the end of the HPCS system life.

All hardware and storage media used for backup shall be provided by the Contractor.

C.4.7.5 Project Management Requirements

The contractor shall be responsible for all functions related to managing the NOAA HPCS R&D system as a single project. NOAA will retain responsibility for project oversight and directing project activities (such as resource allocation) that directly impact the Government. The contractor shall establish communication avenues to keep NOAA managers apprised of daily status and alerts. A formal project management review process shall also be established, as described below.

Reference materials for standard Government Information Technology project management are available on the Internet, for example:

http://cio.doe.gov/ITReform/sqse/publications.htm#Checklists

The contractor shall appoint a Project Manager and assistants as required. This management team will be responsible for meeting NOAA's technical and business requirements (including security), project planning (cost, schedule quality), resource mapping, identifying dependencies and support issues and managing subcontractors. The Project Manager shall hold an Annual Meeting for NOAA management in Silver Spring, MD to discuss project status and plans for the coming year.

The contractor shall establish a Configuration Management (CM) process including a Configuration Management Control Board to include both contractor and NOAA staff. Configuration management plans (at the minimum for network, hardware and software) shall be developed by the contractor and presented to the CM Board for approval. Not less than monthly CM meetings shall be held to review system status and plans for the future. The contractor shall be responsible for preparing an Annual Configuration Management Plan for presentation to NOAA management in Silver Spring, MD. The Configuration Management Control Board will be responsible for oversight of benchmark workstreams, modifications to the benchmarks and for scheduling system time in support of benchmarking.

C.4.8 Profiles

C.4.8.1 Work Stream (WS) Profile

A workstream (WS) is a single instance of end-to-end processing (including pre and post-processing) that represents a set of job streams with similar characteristics within NOAA. NOAA's R&D HPCS Project Team has defined 9 representative workstreams. More complete descriptions of representative workstreams are in Section J. Previous workstream examples are available in section C.7.7 below.

WS1 CM2-ESM – Coupled Earth System Model
WS2 CM2-HR – Coupled High Resolution Global Model
WS3 HIMF – Very High Resolution Ocean Model
WS4 EMTB – Environmental Modeling Test Bed
WS5 CMDC – Climate Model Development and Calibration
WS6 DAD - Data Assimilation Development
WS7 RUC - Rapid Update Cycle
WS8 WRF-EM
WS9 WRF-CHEM

C.4.8.2 Data Generation Profile

The following table shows the projected data generation associated with each workstream throughout the contract period.

Data Generation (TB/day)

	WS1	WS2	WS3	WS4	WS5	WS6	WS7	WS8	WS9
FY2006							2.4	2.2	1.6
FY2007	7	13	5	3	3	3	2.4	2.2	1.6
FY2008	11.2	20.8	8	5	3	5	4.8	4.4	3.2
FY2009	11.2	20.8	8	7	3	7	4.8	4.4	3.2
FY2010	21	39	15	10	6	10	9.5	8	6
FY2011	21	39	15	13	6	13	9.5	8	6
FY2012	35	65	25	20	6	20	12	10	8
FY2013	35	65	25	30	12	30	12	10	8
FY2014	35	65	25	30	12	30	12	10	8

WS1 – WS3 were computed based upon a theoretical proposed 10x computational increase. This increase is based upon the equation of "Compute to the .7 power".

C.4.8.3 Data Retention Profile

The data life-cycle is the expected number of years that the data from a workstream will be stored. These data are all of the remaining files from a given workstream. Including, but is not limited to, source code, input files, run time scripts, generated output, and analysis results.

Data Retention Table

Workstream	Data Life-Cycle
WS1	100% of the data for 9 years; 50% for four score years
WS2	100% of the data for 9 years; 50% for four score years
WS3	100% of the data for 9 years; 50% for four score years
WS4	100% of the data for the first year; 50% of the data for the second year; and 15% of the data past the third year.
WS5	100% of the data for the first year; 50% of the data for the second year; and 15% of the data past the third year.
WS6	100% of the data for the first year; 50% of the data for the second year; and 15% of the data past the third year.
WS7	100% of the data for the first 3 DAYS; 20% of the data for the first year; 10% of the data for 3 years; and 5% of the data for 8 years.
WS8	100% of the data for the first 3 DAYS; 20% of the data for the first year; 10% of the data for 3 years; and 5% of the data for 8 years.
WS9	100% of the data for the first 3 DAYS; 20% of the data for the first year; 10% of the data for 3 years; and 5% of the data for 8 years.

C.4.8.4 Funding Profile

The following table shows the projected funding levels associated with each workstream. The funding begins for workstreams 7-9 in FY2006. The funding for the remaining workstreams will begin in FY2007. The Government requires a reserve option where no less than 94% of the annual funding will be dedicated to the components of the NOAA R&D HPCS specified in the RFP. It is anticipated that funding levels will remain constant for the duration of the base and options periods. Funding shall not cross workstreams, as shown in the table below. Based on information provided by the Contractor, the Government will evaluate proposals in order to verify that workstreams are appropriate to their funding profile.

Funding Profile Table in $Millions (shown before 6% reserve is removed)

	WS1	WS2	WS3	WS4	WS5	WS6	WS7	WS8	WS9
FY2006							$1.3	$1.0	$0.7
FY2007 - FY2009	$4.0	$6.0	$4.0	$2.0	$2.0	$2.0	$1.3	$1.0	$0.7
Option Period FY2010 - FY2013	$4.0	$6.0	$4.0	$2.0	$2.0	$2.0	$1.3	$1.0	$0.7
1 year Options FY2010/ FY2014	$2.0	$3.0	$2.0	$1.0	$1.0	$1.0	$0.7	$0.5	$0.3

C.4.8.5 User Profile

The following shows the percentage of HPC Resources allocated for the representative workstream (WSx) according to the locality of scientific users.

Scientific users do not include Computer Operators, System Engineers, System Administrators, and administrative users.

Users of Workstreams by location (shown in %)

	WS1	WS2	WS3	WS4	WS5	WS6	WS7	WS8	WS9
Boulder				10%	5%	5%	65%	80%	80%
Washington				80%	80%	80%	10%	10%	
Princeton	96%	96%	96%	5%	10%				
Other	4%	4%	4%	5%	5%	15%	25%	10%	20%

User activities may be grouped into the following four categories:

1. Software and Model Development
2. Data Browsing
3. Model Production
4. Model Analysis

Note that a single user may be involved in more than one activity.

Software and model development includes activities such as code development, debugging and optimization as well as the scientific development of the models. The batch jobs submitted by this user set tend to be short, quick turnaround, but resource intensive (large PE counts, large data sets from archive or long term scratch, etc.) consistent with development of next generation models. Use of editors such as XEMACS, the TotalView visual debugger and visualization tools such as grads and ncview characterize many of the interactive tools utilized by this user set. The interactive activities require good bandwidth as described below, but "LAN-like" moment to moment quality of service.

Data browsing is an activity ubiquitous to the user environment. Tools such as grads and ncview are employed to create a graphic representation of spatial and time domain data used as input to the models or created as part of model output. Those engaged in climate workstream activities such as WS1-3 and 5 tend to deal with large data files 1GB to 50GB while weather workstreams such as WS 4,6 and 7-9 with files under 1GB. Such activities require good bandwidth as described below and "LAN-like" moment to moment quality of service. As a further distinction, climate data browsing is generally an "unstructured" activity (i.e. it's not generally known in advance what data will be accessed) while weather data browsing can be viewed as "structured" (i.e. the desired datasets are known in advance).

Model production is characterized by data staging and model runs generally via the submission of batch scripts. This activity includes post processing such as that for WS1-3, 5 and 8 as well as production of graphics files such as that associated with WS7-9. These production runs can be relatively few, medium to long running, self re-submitting jobs such those characterized by WS1-6 or many short jobs such as WS7-9.

Model analysis is characterized by structured and unstructured data browsing and well as interactive and/or batch jobs to produce verification statistics.

C.4.8.5.1 User Profile for WS7 – WS9

NOAA has approximately 200 scientific user accounts associated with these workstreams. Approximately half of those users are at NOAA/FSL. Of the approximately 100 users outside NOAA/FSL over 75% are in the Boulder area. There are approximately 40 users engaged in software and model development, 80 users engaged in data browsing roughly 30% of their time, 15 users engaged in model production and 60 users engaged in model analysis. Data browsing tends to be of the structured type on files 1GB or less. To achieve LAN-like interactive use, each data browsing session requires X Mb/s producing Y images per second to the screen with no perceptible unevenness in image delivery. At time of initial delivery, a minimum of Z simultaneous data browsing sessions shall be supported. Further, a minimum of ZZ interactive TotalView debugger sessions shall be supported with LAN-like interactive use. See Section E for acceptance testing requirements.

C.4.8.5.2 User Profile for WS4 – WS6

NOAA has approximately 200 scientific user accounts associated with these workstreams. Approximately 80% of the users reside in the Washington DC metropolitan area. The remainder is spread across the country with some concentration at the NOAA labs in Princeton, Boulder and Miami. There are approximately 50 users engaged in software and model development, 70 users engaged in data browsing roughly 20% of their time, 10 users engaged in model production and 70 users engaged in model analysis. Data browsing tends to be of the structured type on files 1GB or less for those engaged in WS4 and 6 while WS5 is more unstructured browsing on files greater than 1GB. To achieve LAN-like interactive use, each data browsing session requires X Mb/s producing Y images per second to the screen with no perceptible unevenness in image delivery. At time of initial delivery, a minimum of Z simultaneous data browsing sessions shall be supported. Further, a minimum of ZZ interactive TotalView debugger sessions shall be supported with LAN-like interactive use. See Section E for acceptance testing requirements.

C.4.8.5.3 User Profile for WS1 – WS3

NOAA has approximately 215 active (used in the last 90 days) scientific user accounts associated with these workstreams.133 of the active accounts are users in the Princeton area. There are approximately 30 users engaged in software and model development, 80 users engaged in data browsing roughly 25% of their time, 12 users engaged in model production and 80 users engaged in model analysis. Data browsing tends to be of the unstructured type on files 1GB to 50GB. To achieve LAN-like interactive use, each data browsing session requires X Mb/s producing Y images per second to the screen with no perceptible unevenness in image delivery. At time of initial delivery of resources supporting WS1-3, a minimum of 30Z simultaneous data browsing sessions shall be supported for the Princeton, NJ user set. Further, a minimum of ZZ interactive TotalView debugger sessions shall be supported with LAN-like interactive use for the Princeton, NJ user set. See Section E for acceptance testing requirements.

C.4.9 Security

To assure an adequate level of protection for in-house or commercially maintained IT systems, NOAA maintains all systems consistent with government-wide laws and regulations. The Office of Management and Budget (OMB) Circular A-130 requires all federal agencies to plan for the security of all IT systems throughout their life cycle. OMB Circular A-130 also establishes a minimum set of controls to be included in Federal automated information security programs; assigns Federal agency responsibilities for the security of automated information; and links agency automated information security programs and agency management control systems established in accordance with OMB Circular A-123.

OMB Circular A-130 requires agencies to implement and maintain a program to assure that adequate security is provided for all agency information collected, processed, transmitted, stored, or disseminated in general support systems and major application. It further directs each agency to implement Policies, standards, and procedures which are consistent with government-wide policies, standards, and procedures issued by the Office of Management and Budget (OMB), the Department of Commerce (DOC), the General Services Administration (GSA) and the Office of Personnel Management (OPM).

At a minimum, agency programs shall include the following controls in their general support systems and major applications:

- Assignment of Responsibility for Security.
- System Security Plan consistent with guidance issued by the National Institute of Standards and Technology (NIST) to include:
    - Rules of Behavior/Application Rules.
    - Training.
    - Personnel Controls.
    - Incident Response Capability.
    - Continuity of Support/Contingency Planning.

    - Technical Security/Controls.
    - System Interconnection/Information Sharing.
    - Public Access Controls as required.
- Scheduled Review of Security Controls commensurate with the acceptable level of risk to the system.
- Authorize Processing (Certification and Accreditation).

The Federal Information Security Management Act of 2002 (FISMA), (HR 2548 E-Government Act, TITLE III – INFORMATION SECURITY, SEC. 301. INFORMATION SECURITY), addresses the program management and evaluation aspects of IT security.

Additionally, NOAA follows the US DOC Manual of Security Policies and Procedures (April 4, 2003), the US DOC IT Security Program Policy and Minimum Implementation Standards (July 28, 2004), as well as NOAA OCIO Policies and Directives.

More information can be found at:

- http://www.cio.noaa.gov (NOAA CIO web-site)
- https://www.csp.noaa.gov (NOAA IT Security Office web-site)
- http://www.osec.doc.gov/cio/oipr/ITSec/ITSECDOC1.HTM (DOC IT Security Program Office web-site)
- • http://csrc.nist.gov/policies/ (NIST Federal Requirements)
- • http://csrc.nist.gov/publications/nistpubs/index.html (NIST Guidelines)

In addition, all solutions shall comply with all DOC and NOAA security policies and industry best practices. NIST guidelines/practices are preferred industry best practices.

In order to maintain a consistent security posture across potentially multiple sites, the Contractor shall be required to implement equivalent Patch Management, Change Management, and Perimeter Protection procedures for each sub-System. These procedures shall be documented and maintained as part of the required NIST SP 800-37 Certification and Accreditation package. Although the R&D HPCS will be managed as a single system regardless of physical location, it is expected that multiple sites must rely on public Internet infrastructure for inter-site communication. Special care must be taken to secure intra-site communication over an untrusted network, and mitigate the risks associated with the inherent trust relationships required for full R&D HPCS integration.

C.4.10 Benchmark Performance

Within a workstream's budget, the Government requires the maximum system life throughput utilizing a balanced solution for each workstream in the contract period. The Government considers every workstream to be of equal importance and expects significant performance increases for all of them.

The Government expects a significant performance increase for each workstream to parallel or to exceed Moore's Law. The Government expects that disk and interconnect performance will have increases exceeding those in Moore's Law. The Government also expects security costs will slightly increase but service costs will decrease for this contract compared to the previous ones.

The workstreams were baselined on different computer systems. Namely, WS1 - WS3 were baselined on NOAA's Origin 3000 HPC in Princeton, NJ; WS4 - WS6 were baselined on NOAA's IBM Power 4 Regatta H HPC in Washington, D.C.; and WS7 - WS9 were baselined on NOAA's Xeon-based HPC cluster in Boulder, CO. Further details of the workstream baselines can be found in Section J.

Since the workstreams have different baselines, the Government expects somewhat different initial performance increases. The Government expects a large performance increase for WS1 – WS3, because the baseline was created on a system that was delivered in the spring of 2002 and these workstreams will receive a significant increase in funding prior to FY2007 resulting in computation upgrades. Furthermore, WS4 – WS6 were baselined on computer systems built in 2002 and acquired with substantially higher (99%) availability requirements in summer 2004. The previous contract for WS4 - WS6 also includes significant facilities lease charges, which may either, continue into the next contract or be replaced by wide-area networking charges. Therefore, the Government expects the initial performance increases for WS4 - WS6 to be substantial. Furthermore, WS5 is substantially a pure capacity benchmark whose baseline was run on a capability computer system, so the Government expects a further initial performance increase for WS5.

The Government expects a mid-life upgrade (for example, the second year of a four year contract) for each workstream to achieve performance increases that parallel Moore's Law. The Government can most effectively use performance increases that are implemented temperately during the contract. The Government will have difficulty in effectively using high performance delivered very early in the contract period or very high performance delivered late in the contract period.

The initial deliveries for WS7 - WS9 are required in Q1FY2006. The initial deliveries for WS1 - WS6 are required in Q1FY2007. The mid-contract upgrades for WS7 - WS9 are desired to occur during Q4FY2007 or Q1FY2008. The mid-contract upgrades for WS1 - WS6 are desired to occur in Q2FY2008 or Q3FY2008.

It is worth repeating that the Government desires the maximum System Life Throughput for each workstream in the contract period.

The Government will evaluate performance increases in terms of both capacity and capability as described in Section J. The Government requires maximum System Life Throughput obtained by a significant mid-life upgrade that will not front-load or back-load performance. The mid-contract upgrades for WS1 – WS6 are desired to occur in Q2 or Q3FY2008. The mid-contract upgrades for WS7 – WS9 are desired to occur during Q4FY2007 or Q1FY2008.

C.5 Facilities, GFE

C.5.1 Facilities

The Government does not mandate a facility solution. Rather, the Government requires the Contractor to provide a facility plan for housing the proposed equipment to meet the Government's requirements over the life of the contract. In the plan, Contractors may propose to use any of the following to meet the Government's needs: Contractor-provided facilities; any of the available Government facilities (as described in Section C.8), either as is or with modifications; or a combination of Contractor facilities and Government facilities.

Contractor-provided facilities will have the following conditions:

- Contractors are responsible for providing all facilities resources, including floor space, utilities, facilities maintenance, janitorial services, etc. Therefore, any lost system time caused by environmental outages (such as loss of power, cooling, etc.) or facility failures will be recorded as downtime for availability calculations.
- Contractors may use any Unrestricted GFE provided in Section C.5.2, but are responsible for all shipping costs, both to the Contractor site and back to the Government site at the end of the contract.
- Designated Government personnel must have suitable access to the Contractor facility on an as-needed basis, subject to reasonable vendor restrictions.
- The Contractor's proposal must include a description of the proposed facilities, including:
    1) A copy of the site operating plan
    2) A copy of the security procedures
    3) A guarantee of the maximum number of facility incidents and annual outage time
    4) A statement documenting any single points of failure in computer conditioning
    5) A description of how NOAA remote computer operators will be informed of deteriorating facility conditions such as rising room temperatures or an air handler failure
    6) A copy of the disaster recovery plan
    7) One-line (logic) diagrams of the electrical service and cooling service
    8) An energy density (watts per square foot) projection plotted over the contract life
    9) A spreadsheet listing the type and age of facility equipment to be used. Examples are: UPS systems and power conditioners, chillers, heat exchangers, air handlers.
    10) A copy of the contract statement of work for any commercial facility management company used, or the equivalent if preformed in- house. PM schedules, proactive inspections, and quality assurance methods are examples.

11) A brief (2-3 paragraphs) description of the procedures used to acquire off-site emergency service including minimum response times and escalation procedures
12) A statement (one paragraph) as to how coverage and services are made available after-hours and on holidays
13) A statement (paragraph) projecting the minimum UPS power protection period (survival time) when utility power fails. The worst-case instance covers full functionality of the system(s), and over the entire period of performance (i.e., should load vary during the contract, or battery performance decline due to age).
14) A brief description of the fire protection systems and certification standards
15) A list of design standards complied with by the facility architect (see above)
16) A brief description of any automated facility controls such as computer-managed failover systems
17) A brief description of quality assurance control processes for defect corrections
18) A brief description of the tools used by management to track performance of service level commitments and to manage the facility (Examples: CAD programs, equipment inventory management software, capacity planning and tracking, etc.)
19) A bio (curriculum vitae) of the facility manager's experience and training
20) A description of facility alterations and changes to be made to the offered space if the Offeror is successful
21) Cost data as described above

- The Government reserves the right to conduct site visits of all proposed facilities during the procurement evaluation. During the site visit, the Government may inspect:
    1) Mechanical rooms
    2) Raised floor plenums; drainage; cable tracks, labeling, and management
    3) Network cabling protection and redundancy provisions
    4) Power and cooling distribution and control systems
    5) Logs of equipment failures, corrective actions taken and maintenance results
    6) Preventive maintenance schedules
    7) Emergency plans, and results of drills
    8) Testing procedures and schedules
    9) Training provided to facility managers, maintenance and security personnel
    10) Safety and fire protection equipment and operation
    11) Any other relevant materials that will enable the government to assess the reliability and safety of the facility

In their facilities plan, Contractors may also use the projected facilities resources offered at three NOAA locations, BLDR-1, BLDR-2, and PRTN, which are described briefly below and in more detail in Section C.8. Section C.8.11 indicates the government-supported resources that are projected to be available, beginning at the times indicated and are provided under the assumptions indicated in that section. If Contractors require additional facility resources at these locations beyond what is indicated, they must provide a modification plan indicating how they will obtain these resources. The modification plan must be in accordance with site restrictions.

Contractor use of the government-supported resources implies the following:

- The Government is responsible for providing the indicated facilities resources. Therefore, any lost system time caused by environmental outages or facility failures will be recorded as null time for availability calculations.
- The Contractor is responsible for any facility modifications (e.g., new electrical equipment, plumbing, additional air conditioning equipment) needed to support the proposed systems to be installed. A description of these modifications must be included as part of the Contractor's facility plan.

The Government offers the following facilities for possible use:

Two of the facilities, designated as BLDR-1 and BLDR-2, are located in the David Skaggs Research Center (DSRC), 325 Broadway, Boulder, CO, and the third, designated as PRTN, is located in the Geophysical Fluid Dynamics Laboratory (GFDL) building complex, referred to as the "Princeton Complex" in Section C.8, at 201 Forrestal Road, Princeton University Forrestal Campus, Princeton, NJ.

BLDR-1 currently houses NOAA's JET computer system, which is operated by OAR's Forecast Systems Laboratory under a contract with HPTi. The BLDR-1 facility will not be available until October 2006. BLDR-2 is a facility currently under construction at DSRC and will be available for use in October 2005. PRTN currently houses NOAA's HPCS, which is operated by OAR's Geophysical Fluid Dynamics Laboratory under a contract with Raytheon Company.

Appendix B (Section C.8) provides descriptions of these government-furnished facilities, which are designed to support high-performance computer systems. These sites are available for possible use by Contractors to house systems under the NOAA HPC R&D contract, subject to the availability of facility resources as projected in Section C.8.11 under the assumptions provided therein

The following table summarizes projected raised floor space by location and earliest date of availability:

Projected Resources	Site	October 2005	October 2006
Raised Floor Space (sq.ft.)	BLDR-1	0	2250
	BLDR-2	1500	1500
	PRTN	2700	5500

The Government will offer site visits to Offerors on the following schedule:

BLDR-1 and BLDR-2 ....January 18, 2005
PRTN ....January 20, 2005

During the site visit, Offerors may obtain a set of blueprints for that facility once an authorized representative signs a non-disclosure agreement and agrees to destroy all blueprints at the end of the solicitation. Note in particular that Offerors requesting building diagrams or related documents of federal government facilities (specifically all documents relating to DSRC) shall be required to have an authorized representative execute a non-disclosure form as required under GSA PBA-3490.1.

C.5.1.1 Procedures for government-designated personnel to access contracted facilities

If the contractor proposes the use of non-Government facilities, he must provide procedures that allow government-designated personnel to access these facilities.

C.5.1.2 Government approval of construction timing and impacts

C.5.1.3 Policies for providing workstations, phone service, office equipment to on-site contract personnel

BLDR-1 & BLDR-2
Current policy is one where the government provides phones and LAN connections via Standard Information Outlets (SIO) only. All workstations and office/lab furniture will need to be supplied by contract personnel.

PRTN
PRTN’s policy under the current Raytheon contract is to require the vendor to provide phone service, workstations, and office furniture for Computer Engineers and Computer System Administrators located in the Computer Building. However, vendor-supplied software analysts, security specialists, and network support personnel use government-supplied workstations, office furniture, and phone service, because they are typically integrated into the laboratory’s normal office infrastructure, rather than being located in offices in the Computer Building that have been designated for vendor system support personnel.

C.5.1.4 Procedures for obtaining construction permits

BLDR-1 & BLDR-2
Construction permits are not required because DSRC is a federal building.

PRTN
Under the current contract, any construction to be done requires the Contractor to obtain construction licenses, permits, and approvals for this work from Plainsboro Township and Princeton University, as required.

C.5.1.5 Costs for facilities modifications in GSA properties

GSA Regulations require that GSA contractors carry out all modifications to GSA buildings. Therefore, any facilities modifications will require that funds be held back from the contract for transfer to GSA for any changes to be made to GSA properties.

BLDR-1 & BLDR-2
Close coordination with the DSRC landlord, U.S. General Services Administration (GSA), will be required if any building modifications are required. Costs will include normal GSA overhead charges plus project management oversight. All modifications must meet standard building codes as dictated by GSA. Modifications can include small projects like power receptacle changes, to large projects like chiller plant upgrades. Normal government contracting procedures can be expected for any modifications.

PRTN
Not applicable

C.5.2 Government Furnished Equipment (GFE)

The following describes Government-owned property that is available to be furnished to the Contractor for the performance of this contract. Because some Government-furnished equipment that is located at the designated sites is only available for use by the Contractor if that facility is used, the GFE lists are separated into two categories: “Site-Constrained GFE”, which is equipment that is only available if used at the designated site; and “Unrestricted GFE”, which is available for use under this solicitation without site restrictions.

C.5.2.1 Boulder, CO

Site-Constrained GFE

- Partial use of UPS and Cooling equipment in BLDR-1. The equipment is maintained by the Government (250kVA and commensurate cooling).
- Partial use of UPS and Cooling in BLDR-2 (350kVA and commensurate cooling). The equipment is maintained by the Government.
- XXX Myrinet Y-Cards, ZZZ Myrinet Z-Cards, QQQ Myrinet switches (available 10/2006)
- An XXX Extreme Networks gigabit ethernet switch (available 10/2006)
- 2 DDN S2A6000's with (available 10/2006)
    - XXX drives
- YYY drives

Unrestricted GFE

- None

C.5.2.2 Washington, DC

Site-Constrained GFE

- None

Unrestricted GFE

- None

C.5.2.3 Princeton, NJ

<u>Site-Constrained GFE</u>

Manufacturer	Part Number	Description	Qty
		Computer Room Security Monitoring	
		C-Cure 800	
Compac	4403US	Compac MS 2000 C-Cure 800 OS	1
Emerson	P761VBT0EENC	Monitor C-Cure 800	1
		Computer Room Security Monitoring	
		Intellex 16000	
Sensormatic	DVMS DV16000	Intellex 16000 Video recorder system w/CD backup max 16 cameras	1
NEC	FE1250+BK	Monitor Intellex 16000	1
Pelco	CC-3700-S	Color CCD Camera	12
		Power Distribution Units	
United Power	PDM4-F3-225-K13-426	PDU 1 (225 kVA)	1
United Power	PDM4-F3-225-K13-426	PDU 2 (225 kVA)	1
United Power	PDM4-F3-225-K13-426	PDU 3 (225 kVA)	1
		Uninterruptible Power Supply	
MGE	72-131522-000	UPS Cabinet and Battery Bank 225 KVA	1
		Power Distribution Units	
Liebert	PPA125C	PDU Unit (125 KVA)	1
EPE	PD084M48A12-125	PDU Unit (125 KVA)	1
		Computer Room Air Conditioners	
DataFlow	CCT-60C4	CRAC 1 (35 Tons)	1
DataFlow	CCT-60C4	CRAC 2 (35 Tons)	1
DataFlow	CCT-60C4	CRAC 3 (35 Tons)	1
DataFlow	CCT-60C4	CRAC 4 (35 Tons)	1
DataFlow	CCT-60C4	CRAC 5 (35 Tons)	1
DataFlow	CCT-60C4	CRAC 7 (35 Tons)	1
APC	CCT-60C4	CRAC 8 (35 Tons)	1
APC	CM-3.0-W-BC-D	CRAC 6 (3 Ton)	1

Unrestricted GFE

Manufacturer	Part Number	Description	Qty
		Switch for Shared Storage	
SGI	FC-SWITCH-16	SGI Fibre Channel switch with 16 ports and one power supply	8
SGI	FC-SWITCH-PWR	Optional second power supply for SGI FC switches (8 & 16 port)	8
SGI	FC-SWKIT	Rackmount kit for mounting FC-SWITCH-8 or FC_SWITCH-16 in F1RACK	8
SGI	XSWOPTGBIC	Short Wave Optical GBIC kit containing 6 GBICs	16
SGI	XCOPGBIC	Copper GBIC kit containing 6 GBICs	6
		Switch for Tape Drives	
SGI	FC-SWITCH-16	SGI Fibre Channel switch with 16 ports and one power supply	8
SGI	FC-SWITCH-PWR	Optional second power supply for SGI FC switches (8 & 16 port)	8
SGI	FC-SWKIT	Rackmount kit for mounting FC-SWITCH-8 or FC_SWITCH-16 in F1RACK	8
SGI	XSWOPTGBIC	Short Wave Optical GBIC kit containing 6 GBICs	16
SGI	XCOPGBIC	Copper GBIC kit containing 6 GBICs	6
		Console and Monitoring	
SGI	SG230-00008	230L Workstation, 667 MHz PIII, 128MB PC 133 SDRAM, 20GB IDE, V3 Gfx 32MB DDR, Red Hat 6.1	2
SGI	91-AB945-001	19" Northern Hemisphere Monitor	2
SGI	91-AD001-001	Keyboard, Mouse, Speakers, Power Cords, Monitor Cable, User Manual	2
SGI	SSU60003	HW on-site support 4hr rsponse, 7x24, years 1-3 for SGI 230 Workstation	2
SGI	SC4-PCP-2.0	Performance Co-Pilot - Performance Monitoring tool for IRIX 5.3 and higher (replaces PCPORIGIN)	1
SGI	SC4-PCPHPC-1.0	Performance C-Pilot Add-On agent for IRIX 6.5 clusters (replaces SC4-PCPARRAY-1.0); requires SC4-PCP-2.0	8
SGI	SV4-PCPCOL-2.0-10	Performance Co-Pilot Collector 10 license pack	1
		Hierarchical Storage Mgmt System	
SGI	ORIGIN-3800	SGI Origin 3800 server - 64 CPUs (GFE) 600Mhz R14000A processors), 64GB memory, 16 Local Disk Channels (1Gb/sec), 16 Shared Disk Channels(2Gb/sec), 24 Tape I/O Channels, and 2 Gigabit Ethernet Channels. 180GB	1

StorageTek	9310002-0000	POWDERHORN 6000 CART/450 EPH	3
StorageTek	9940L03-0000	9940, Library, Fibre	24
StorageTek	9840L03-0000	9840, Library, Fibre	22
SGI	FC-SWITCH-16	16 port 1Gb FC switch (8Cu/8Optical)	12
SGI	ORIGIN-3800	SGI Origin 3800 server - 64 CPUs (Lease 1a) 600Mhz R14000A processors), 64GB memory, 16 Local Disk Channels (1Gb/sec), 16 Shared Disk Channels(2Gb/sec), 24 Tape I/O Channels, and 2 Gigabit Ethernet Channels.	1
	Storage	HSMS Local Storage	
	SGI TP9100 1GB	TP9100 D-Brick with fourteen 36GB 10KRPM Drives	16
	STOR-CTRL 128	TP9100A Dual Channel Control Unit with 1Gb FC	16
	SGI TP9100 1GB	TP9100 D-Brick with fourteen 18GB 10KRPM Drives	2
	STOR-CTRL 128	TP9100A Dual Channel Control Unit with 1Gb FC	2
	Empty Rack	I/O Racks with AC Power Distribution	3
	FC-SWITCH-16	16 port 1Gb FC switch (8Cu/8Optical)	4

		Hierarchical Storage Mgmt System	
StorageTek	9310002-0000	POWDERHORN 6000 CART/450 EPH	1
StorageTek	9840	Media Cartridges	3000
StorageTek	9940	Media Cartridges	12000
	Connectivity	Connectivity - Existing	
Cisco	WS-C6509	Catalyst 6509 Chassis	1
Cisco	WS-CAC-1300W	Catalyst 6000 1300W AC Power Supply	1
Cisco	WS-CAC-1300W/2	Catalyst 6000 Second 1300W AC Power Supply	1
Cisco	SFC5K-SUP-5.5.1	Catalyst 6K Supervisor Flash Image, Release 5.5(1)	1
Cisco	WS-X6K-SUP1A-2GE	Catalyst 6000 Supervisor Engine 1A, Enhanced QoS, 2GE	2
Cisco	WS-X6416-GE-MT	Catalyst 56000 16-port Gig-Ethernet mod. MT-RJ	2
Cisco	WS-X6408A-GBIC	Catalyst 6000 8-port GE, Enhanced QoS	2

Cisco	WS-X6348-RJ-45	Catalyst 6000 48-port 10/100, Upgradable to Voice	2
Cisco	WS-G5486	1000BASE-LX/LH "long haul" GBIC	4
Cisco	WS-G5484	1000BASE-SX 'Short Wavelength' GBIC (Multimode only)	36
Cisco	CISCO3640	Cisco 3600 4-slot Modular Router-AC with IP software	2
Cisco	S364C-12.0.4T	IP	2
Cisco	NM-2CT1-CSU	2-Port Channelized T1/ISDN-PRI with CSU Network Module	2
Cisco	NM-1FE1CT1-CSU	1 Port F Ethernet 1 Port Channelized T1/ISDN-PRI with CSU NM	2

		Software	
Gridware	S8102	GRD License for a single CPU, CODINE included. Software is binaries and licenses delivered via Internet. No media is provided. Optional product manuals sold separately.	1
Gridware	S8201	Codine one year support for a single CODINE license.	1
Etnus	TV-SGI-4u-128p	TotalView	2
Etnus	TV-SGI-4u-256p	TotalView	1
Mathworks	Matlab	Matlab	10
Wolfram Research	Mathematica	Mathematica 4.0 Network licenses	10
Mathsoft	S-Plus	S-Plus 5.1 Unix SGI	10
NCAR	NCAR	NCAR	1
Legado	SR4-SILO128-5.5	Silo Software Module for Networker with 1-128 managed volumes	1
SGI	SC4-NWKRPE-5.5	Networker 5.5 Power Edition, Supports up to 10 clients, 64 simultaneous data streams & 32 de	1
SGI	SC4-NWKRPE-5.5	Networker 5.5 Power Edition - Maintenance	1

C.6 Options

C.6.1 Option Period

The Base Period of the contract is FY2006-FY2009. The Government requires an option to extend the contract for another four years from FY2010-FY2013. Requirements are approximately listed through Section C.

C.6.2 One-year extension of Base Period

After the four-year Base Period, the Government requires the option to extend the lease for operations, maintenance and the equipment for one additional year in quarterly increments. If exercised, funding is shown as the FY2010 level as expressed in C.4.8.4.

C.6.3 One-year extension of Option Period

After the four-year Option Period, the Government requires the option to extend the lease for operations, maintenance and the equipment for one additional year in quarterly increments. If exercised, funding is shown as the FY2014 level as expressed in C.4.8.4.

C.6.4 Additional R&D HPCS Augmentations

Over the life of the contract there may arise situations that will require the R&D HPCS to be augmented. One such situation might result when NOAA identifies a new requirement for HPC that did not exist at the time of contract award. A second situation that might arise would be if NOAA was to enter into an inter-agency agreement with another Federal agency to supply that agency with computational resources. In either situation the Government shall request the Contractor to provide a proposal to meet any such requirement.

C.6.5 Engineering Support

The Government requires additional expert-level engineering to address impending needs for a given workstream.

C.6.5.1 Applications Analyst

C.6.5.2 Systems/Network/Security Engineer

C.6.5.3 Facilities Engineer

NOAA may require professional engineering services in support of it's High Performance Computing (HPC) facilities. These services may be architectural, electrical, mechanical or civil engineering services. Activities may include, but are not limited to the following:

1. Facility Engineering Consulting Services
2. General Engineering Studies
3. Specific Engineering Studies
4. Design Services
5. Design-Build Services

Engineer Qualifications:
All services will need to be performed by engineers who meet the following minimum requirements:

1. 10 years experience in High Performance Computer (HPC) facility engineering.
2. Past history of successful HPC facility engineering services.

3. Specific HPC power, cooling and infrastructure experience.
4. Membership in good standing in appropriate engineering trade organizations.
5. Past performance in Federal Government contracting.

The Government desires the following Additional qualifications:
1. Professional Engineering (PE) services.
2. Engineering Stamps.
3. Plan draft, review and approval.
4. Drawing update and certification.

Period of Performance:
Engineering services may be required on a daily, weekly, monthly or quarterly basis.

C.7 Appendix A – Details of current NOAA R&D HPCS

C.7.1 Large-Scale Computing (LSC)

C.7.1.1 Boulder, CO

NOAA's Large Scale Computing located in Boulder consists of five logical clusters based on Intel 2.2GHz Xeon systems interconnected with MyriNet. Each node has two CPUs and 1 GB of memory. Two clusters contain 62 nodes each, one cluster contains 128 nodes, and two clusters contain 256 nodes. There are four front-ends with 4 GB each. There is also a test bed containing 12 dual-processor Opteron systems and 12 dual-processor Itanium systems.

C.7.1.2 Washington, DC

NOAA's Large Scale Computing located in Washington, DC consists of 40 IBM Power4 1.3 GHz Regatta H compute nodes, with a total of 1280 processors. Each node has 8 LPARs with 4 processors per LPAR. Each processor has 1 GB of memory. The interconnect is an IBM SP Switch-2, which provides for multiple plain support for the switching fabric. Each LPAR has 2 SP Switch-2 PCI adapters.

The storage and I/O system is integrated with the switch and contains 4 additional dedicated Regatta H nodes each with 32 processors, configured in 4 LPARs with 8 processors per node. Each processor has 2 GB of memory. Each LPAR has two IBM SP Switch-2 adapters.

C.7.1.3 Princeton, NJ

NOAA's Large Scale Computing located in Princeton is designed for the batch processing of computationally intensive jobs. The LSC consists of 11 nodes with 3040 processors. Fast scratch disk is available from each node. The fast scratch disk has eight 1 Gb FC per host and is in a RAID 5+1 configuration.

There are eight SGI Origin 3000s running Irix 6.5.19f and consisting of:

2 hosts with	512 x 600 Mhz MIPS, 512 GB memory, 2.6 TB fast-scratch disk
5 hosts with	256 x 600 Mhz MIPS, 256 GB memory, 0.6 TB fast-scratch disk
1 host with	128 x 600 Mhz MIPS, 128 GB memory, 0.6 TB fast-scratch disk

There are three SGI Altix 3700s running Red Hat Enterprise AS with SGI Propack and consisting of:

2 hosts with	256 x 1.5 Ghz Intel, 512 GB memory, 1.4 TB fast-scratch disk
1 host with	96 x 1.5 Ghz Intel, 192 GB memory, 2 TB fast-scratch disk

In April 2005, NOAA will receive a performance increment of at least 1.8X above the Origin portion of the Large Scale Computing component.

C.7.2 Post-Processing and Analysis

C.7.2.1 Boulder, CO

There are four visualization servers they are based on Intel 1.7 GHz Xeon systems. They have dual ethernet connectivity (to the internal cluster network and to the internet via NOAA/FSL's LAN). There are also four cron servers based on 2.2 GHz Xeon systems.

C.7.2.2 Washington, DC

NOAA’s HPCS system located at Washington DC contains 16 LPARs that are dedicated to single treaded batch jobs (mostly pre- and post-processing) and for interactive access to the parallel compute environment, which consists of the remainder of this HPCS. Additional post-processing and visualization is performed at several branch servers at NCEP, which are typically Origin 300 or Altix servers.

C.7.2.3 Princeton, NJ

NOAA’s Analysis system located at Princeton has two nodes for analysis, post-processing and development work. The Analysis system is comprised of two SGI Origin 3900 systems and one SGI Onyx 3000 with visualization tools.

2 hosts with	96 x 600 Mhz MIPS, 96 GB memory, 3.9 TB fast-scratch disk
1 host with	4 x 400 Mhz MIPS, 4 GB memory, 0.2 TB scratch disk

The fast scratch disk is comprised of 48 x 1 Gb FC per host that connects to a XFS filesystem in a RAID 5+1 configuration.

The analysis computers are available for interactive use at all time. Parallel applications on up to 16 CPUs may be run for up to 8 hours. For applications that use more than 16 CPUs, use of the Large-Scale Computing is recommended, not required.

C.7.3 Storage and Archiving

C.7.3.1 Home File Systems

C.7.3.1.1 Boulder, CO

There are two Data Direct Networks S2A 6000s each has 8 x 1 Gb/s FibreChannel ports. One has 9.2 TBytes of useable space and the other has 6.7 TBytes of useable space (this space excludes parity stripes). Files systems are served by a number of Linux-based NFS servers with read/write data rates of 40-50 MBytes/second each. There are six Dell 2650s with dual 1 GHz Pentium IIIs, one Dell 2550 with dual 933MHz Pentium IIIs, three dual-processor 2.2GHz Xeon systems, and one dual-processor 2.6GHz Xeon system. All of the systems have 2 GBytes memory, one Gigabit ethernet interface and one 1 Gb/s FiberChannel HBA. These file systems provide long term (but not backed up) data storage. These file systems also provide the /home file system for the clusters.

There is a 1.6TByte scratch file system based upon PVFS that utilizes commodity (ATA) disks for very short-term storage. This file system currently sustains around 250MB/s.

C.7.3.1.2 Washington, DC

The home file systems on NOAA's HPCS system located at Washington DC consist of 25.6 TB of disk space. Most of this disk space is configured as a General Parallel File System (GPFS), which is integrated with the switch as described in section C.7.1.2 above. GPFS utilizes 2 GB Fibre Channel adapters per storage and I/O LPAR.

C.7.3.1.3 Princeton, NJ

In Princeton, NOAA is using a unified /home filesystem for all scientific computers and workstations. This allows for users to log into various computational resources and have their environment and files follow them. This has required custom .cshrc and .login dotfiles that are designed for the unified home directory.

By default, each scientific user is limited to 10 GB of home directory disk space. /home is used primarily for model source code, batch job printouts, and files for workstation applications. Large data files are kept in /archive (the filesystem that is under HSM).

The home file system is a 2.3TB SAN from a SGI Origin 3000 server. This file system is being served via CXFS within the HPCS and via NFS to the scientific workstations. Currently, the filesystem is 60% used by 10 million files. Backups are currently being done via xfsdump to the STK 9940B tapes in the silos.

C.7.3.2 Hierarchical Storage Management (HSM)

C.7.3.2.1 Boulder, CO

NOAA's HSM in Boulder is based upon ADIC's StoreNext software. The robotic component is an ADIC AML/J with 8 LTO tape drives. Managed file systems (disk cache) use the DDN S2A 6000s mentioned above. Currently, 1 TB of space is allocated for two managed file systems. Over 88 TB is managed and is composed of over 4,000,000 files. Access to the HSMS is primarily via locally written get/put scripts, some limited direct access to the managed file systems is allowed.

C.7.3.2.2 Washington, DC

NOAA's HPCS in Washington DC has access to a 2 PB HSM. The HSM is presently considered part of the operational system at NCEP. This storage system, in principle, consists of tape storage with dedicated disk storage for staging data to be stored to, or retrieved from tape. This system is based on HPSS. Access to this system is through htar and PFTP commands, and a public domain HSI Unix-like interface. Additional information on HPSS and IHSI can be found at:

http://www.hpss-collaboration.org
http://www.sdsc.edu/Storage/ihsi/Welcome.html

C.7.3.2.3 Princeton, NJ

NOAA's HSM in Princeton uses SGI's Data Migration software to provide a data archive via the /archive filesystem. This software manages both a disk and a tape copy of each file in /archive. This is done transparently to the user, under one file name.

The HSM software has a 15.8 TB SAN for file staging. The disk SAN is connected to the rest of the HPCS by sixteen 2 Gb FC per node. There is currently 2.5dPB stored in 10.7 million files on 9940B (200GB/tape) and 9840 (20dGB/tape) tapes. There are also 2.4 million files disk-resident files that are smaller than 64 KB.

To archive a file, a user simply copies or moves the file into their /archive/<user> directory from any node in the HPCS. Once a new file is present in /archive, data migration will automatically make a tape copy of the file. At first, the disk copy of the file is also kept available. If the file is not accessed for several days, data migration will remove the disk copy of the file, keeping only the tape copy. The next time this file is accessed, it will be "staged" from tape to disk. Subsequent accesses will then be from the disk-resident copy of the file. However, files may be moved between subdirectories, renamed, or removed without causing the file to be "staged" from tape to disk.

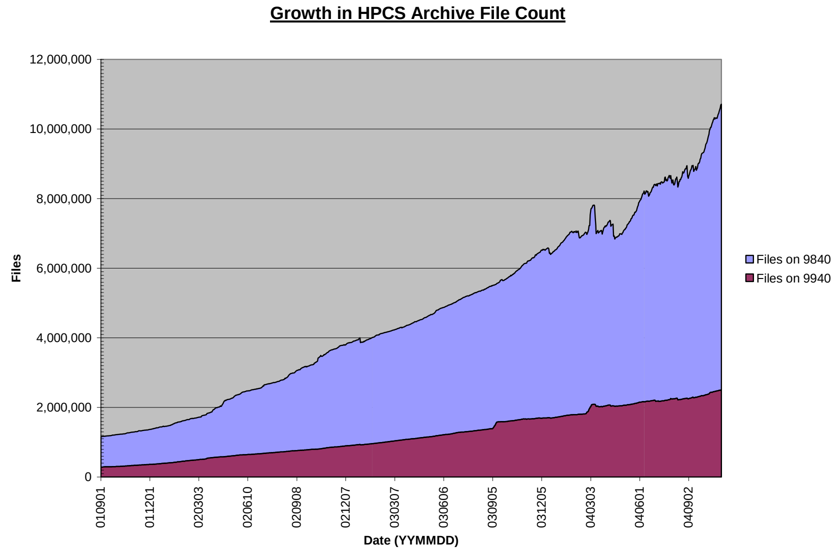

Figure 1 - Growth in tape-resident files on HSMS in Princeton

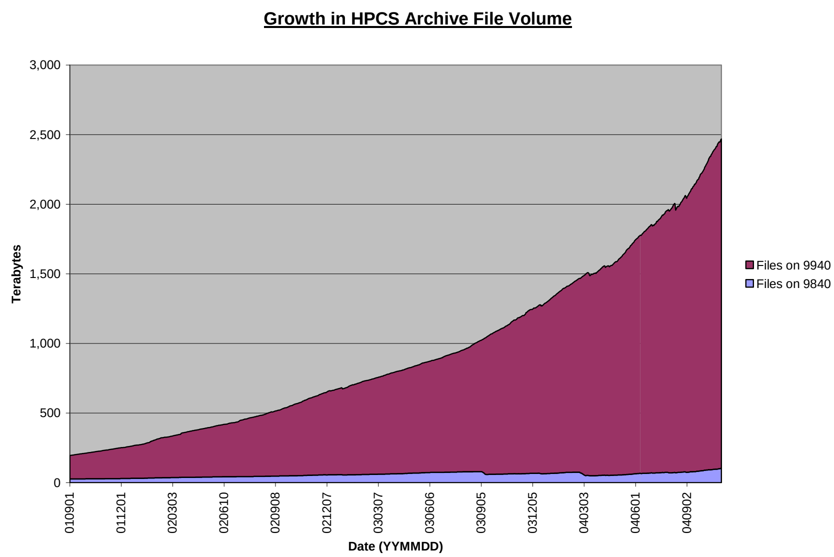

Figure 2 - Growth in tape resident storage on HSMS in Princeton

C.7.4 System Integration

C.7.4.1 Wide Area Networks (WANs)

	Internet	Internet2	Other
Boulder	155 Mb/s	310 Mb/s	NLR is expected Q3FY2005
Washington, DC	622 Mb/s	155 Mb/s	
Princeton	9 Mb/s	100 Mb/s	1Gb/s – Princeton University Sayre Hall (Forrestal campus)

C.7.4.1.1 Boulder, CO

NOAA's Boulder facility is connected to the Abilene network via the Front Range GigaPOP at 310Mb/s. A 155Mb/s connection is available to Boulder through a commercial Internet provider. By the third quarter of Fiscal Year 2005, Boulder will be part of the National Lambda Rail (NLR) with connectivity of at least 1000Mb/s.

C.7.4.1.2 Washington, DC

NOAA's Camp Spring, MD facility, is connected to Internet2 with an OC3. A OC12 internet link is presently used close to capacity. There is some capacity left on the OC3 link. This setup is propagated from the previous operational use of the present HPCS in Washington, DC. It is expected that future HPCS systems will rely more on Internet2.

C.7.4.1.3 Princeton, NJ

NOAA's Princeton facility is connected to Internet2 through Princeton University at 100Mb/s. NOAA's connection to Princeton University is via a 100Mb/s microwave link. A 9Mb/s is available to Princeton through a commercial Internet provider.

A Cisco 2600 router and a Juniper M7i router are used to connect to the commodity Internet, Princeton University, and subsequently Internet2. These exterior border routers, along with the local DMZ networks, are segmented off from the LAN by redundant Cisco PIX 535 firewalls.

C.7.4.2 Local Area Networks (LANs)

C.7.4.2.1 Boulder, CO

NOAA's Boulder LAN backbone that connects to the HPCS is based upon Gigabit Ethernet.

C.7.4.2.2 Washington, DC

NOAA's Washington, DC LAN that connects to the HPCS consists of Cisco Catalyst 3500 XL routers at 100 Mb/s. Apart from several hundred desktop systems, the LAN includes several servers. Servers are typically high-end SGI Origin 300 or SGI Altix servers which are used for storage, post-processing and development.

C.7.4.2.3 Princeton, NJ

NOAA's Princeton LAN backbone that connects to the HPCS consists of Brocade, Cisco, Enterasys and Juniper equipment. The HPC backbone ethernet network is handled by a Cisco Catalyst 6509. The HPC nodes are connected at both 100 Mb/s and 1Gb/s. To support the large data requirements of the system, Brocade Silkworm and Cisco Catalyst switches are used to connect the SAN infrastructure to the HPC nodes. Local users connect to the LAN via distribution switches at either 100Mb/s or 1Gb/s.

Basic network services provided include: DNS, LDAP, Mail, NFS, NIS, NTP, and Printing.

C.7.4.3 Storage Area Networks (SANs)

C.7.4.3.1 Boulder, CO

NOAA's SAN in Boulder is described in section C.7.3.1.1.

C.7.4.3.2 Washington, DC

NOAA's SAN in Washington, DC is described in section C.7.3.1.2.

C.7.4.3.3 Princeton, NJ

NOAA's I/O for its SAN in Princeton is typically dominated by reading restart files at the beginning of a run, writing snapshots of geophysical variables throughout the run (history files), and writing restart data at the end of the run. A restart file is usually a single large binary file, while the history files can be either one large file or many small files (for example, one file written per CPU, typically in netCDF format) that can be merged into a single large netCDF file after the production run finishes. The I/O to large restart files and the efficient creation of many small files on the LSC disk used for temporary storage, as well as the ability of the HSMS (see C.7.3.2) to archive many small files simultaneously are among GFDL's principal concerns.

NOAA's SAN in Princeton has 22.5TB of capacity that is accessible through switched 2 Gb/s FC. The LSC in Princeton has eight 2 Gb/s channels per node to this SAN. The AC in Princeton has twenty-four 2 Gb/s channels per node to this SAN.

C.7.4.4 Batch queuing systems

C.7.4.4.1 Boulder, CO

NOAA's HPCS in Boulder uses the open source version of Sun Grid Engine (SGE). Currently version 5 is in use on the primary set of clusters. Version 6 will be in use by November 2004. With version 5, a custom prescheduler is used to manage jobs based upon the account specified by the user. Accounts are given a maximum number of CPUs to be in use at any given point in time, a maximum number of jobs in use at any given point in time, a maximum job length and maximum priority. Certain accounts are also given dedicated resources (this will be replaced by resource reservations with version 6). Accounts are members of classes (classes can be members of classes as well), classes have total resource limits as well that the prescheduler takes into account. The qsub command is "wrapped" with a custom script that verifies that the user is a member of the account and hasn't exceeded any of the per job maximums. The “wrapper” also modifies priority 0 jobs to use an alternate backfill account associated with the primary account. Typically at least 70,000 jobs per week pass through the batch system.

C.7.4.4.2 Washington, DC

NOAA’s HPCS in Washington, DC uses LoadLeveler. The system is set up with separate partitions for serial and parallel jobs (see section C.7.2.2 above). Queues have been assigned to provide access to specific job resource requirements and priority. Accounting is presently performed for system monitoring only, but is deemed essential is future R&D resource allocation.

C.7.4.4.3 Princeton, NJ

NOAA’s HPCS in Princeton uses Grid Engine Enterprise Edition 5.3. This is a POSIX 1003.2d batch environment (with extensions) that has been provided by Raytheon. IRIX cpuset support has been added. More information on the batch queuing system at http://gridengine.sunsource.net/

In order to run a batch job, a user must prepare a job script in a file, then submit the script to the batch queuing software with the “qsub” command. By default, all jobs are executed by the C shell and must specify a “parallel environment” and a “hard CPU time limit”. The maximum allowed “hard CPU time limit” is 8 hours. The batch workload is overseen by Government operations staff.

Each calendar month, a running total is kept of each group’s CPU usage. Each group is assigned a monthly CPU allocation by the GFDL Director. When a group has exceeded its monthly allocation, that group’s jobs are blocked from running, however the group is still allowed to run jobs in the windfall queue.

Grid Engine projects are used to track usage and assign higher job scheduling priority to some activities. Only users designated by their team may use these projects.

CPU accounting runs once per day, which limits the granularity for enforcing the monthly allocations.

C.7.5 Software

C.7.5.1 COTS

C.7.5.1.1 Boulder, CO

NOAA's HPCS in Boulder currently has licenses to use IDL on the visualization servers. The HPCS in Boulder also has licenses for the Portland Group Fortran and C compilers as well as the Intel Fortran and C compilers.

C.7.5.1.2 Washington, DC

NOAA’s HPCS system located at Washington DC presently has licenses to use the following COTS software on the LSC: TotalView, Vampir, IDL, Matlab. Also available are the following libraries: Visual Numerics IMSL, ScaLaPack.

C.7.5.1.3 Princeton, NJ

NOAA’s HPCS in Princeton currently has licenses to use the following COTS software on the HPCS: CASEVision/WorkShop, MIPSpro f90 Compiler, Etnus TotalView, F-lint, IDL, Legato Networker, Maple, Mathematica, Matlab, MIPSpro f90, NAG Fortran Libraries, NAG Iris Explorer, and S-Plus.

C.7.5.2 Community Supported Software

C.7.5.2.1 Boulder, CO

NOAA's HPCS in Boulder currently utilizes the following open source software: SUSE Linux, Fedora Linux, Ferret, NCAR graphics, GrADS, and NetCDF utilities.

C.7.5.2.2 Washington, DC

NOAA’s HPCS system located at Washington DC presently utilized the following open source software: GrADS, NCAR graphics, ImageMagick, NetCDF utilities, GEMPAK, and ISMF.

C.7.5.2.3 Princeton, NJ

NOAA’s HPCS in Princeton currently utilizes the following open source software: Ferret, Grace, GrADS, NCAR graphics, and NetCDF utilities.

C.7.6 Reliability, availability and support

C.7.6.1 Boulder, CO

The current contractual uptime requirement for the HPCS in Boulder is being exceeded significantly by the incumbent. For carefully monitored jobs, the reliability of the system in terms of delivery of results to an end user is over 98.5%.

The current facility has 250kVA available to the current HPCS and commensurate cooling.

The incumbent supplies two on-site software engineers to support the facility. A program manager is on-site over 50% of the time. Other software engineers work remotely on the system a small (<10%) percentage of time to provide additional support.

C.7.6.2 Washington, DC

Reliability:

Reliability is expressed in terms of Mean-Time-Between-Failure (MTBF). A failure is defined as any scheduled or unscheduled event that triggers down time. Individual components of the system can fail and not be counted as a system failure if the system has sufficient redundancy, error correction or resiliency to allow the failure to not affect individual jobs running in production or development.

Availability:

Availability is determined by computing the ratio of total batch computation processor hours available for execution of the operational and developmental suites to the total batch computation processor hours each month. Time a batch computation processor is available for execution is determined by subtracting processor down time from wall-clock time. Spare nodes or processors may be configured into the batch computation processor pool upon failure of a node and/or processor in the batch computation pool to reduce down time, but down time accumulates until the spare is made available in the system for execution. Failure(s) of other components of the HPCS that result in batch computation processor(s) being unable to run jobs is equivalent to failure of those batch computation processors themselves.

Down Time:

Down time is that period of time when the HPCS or a component thereof is inoperative. Down time shall commence at the time when NOAA contacts the Vendor's maintenance representative at the designated point of contact or the Vendor's answering service or other continuous telephone coverage provided to permit NOAA to make such contact to report a failure. Down time shall end when the HPCS is returned to NOAA in operable condition. The HPCS, or individual component thereof, may be declared inoperative while problem diagnosis takes place. Down time includes any time required for operating software regeneration, reinstallation, or reconfiguration. During a period of down time, NOAA may continue to use operable components of the HPCS when such action does not interfere with maintenance of the inoperable components of the HPCS.

Support:

The Government requires 7x24x365 System Administration support. Current staffing levels are at 3 System Administrators, 1.25 Application analysts and one part time HSM specialist.

C.7.6.3 Princeton, NJ

The HPC System continues the Government's historically high utilization of its computing resources. Reliability, availability, and Contractor support are considered fundamental aspects of the HPCS

Downtime

Downtime is that period of time when all of an HPCS component's workload cannot be accomplished due to a malfunction in the Contractor-maintained equipment or software, or when the HPCS or a component of the HPCS is released to the Contractor for maintenance. Periods of Remedial and Preventative Maintenance count as downtime.

Null time is that period of time when the workload cannot be accomplished due to circumstances beyond the scope of the contract. Null time does not count as downtime.

Downtime for each HPCS component is based on the fraction of the resources available for that component's workload, arrived at through consultation between the Government and the Contractor, and determined ultimately by the Government. Downtime is accumulated on the LSC, Post-Processing and Analysis computers if they are not able to perform the workload when either the HSMS or HFS is down.

The Contractor provides the Government a designated point of contact to request maintenance. The Contractor maintains escalation procedures that allow the Government round-the-clock telephone contact with knowledgeable Contractor staff should the designated point of contact be unavailable.

A component's downtime shall commence at the time the Government makes a bona fide attempt to contact the Contractor at the designated point of contact. At this time the Government begins a log of the problem. Information to be entered into the log is determined by the Government.

A component's downtime excludes any time in which the Government denies the Contractor maintenance personnel access to the malfunctioning hardware and/or software.

A component's downtime ends when the computer is returned to the Government in operable condition as determined by the Government, ready to perform all of the workload.

Preventative Maintenance is completed outside of primetime hours (7 am – 7 pm).

The testing and installation of every major operating system release installed at the request of the Government and one (1) minor operating system release installed at the request of the Government during any annual period will count as downtime.

Preparation for the execution of post-upgrade LTDs, including any benchmark development by the Contractor, associated with the agreed-upon upgrades proposed at contract award, will count as downtime.

Availability

Proposed throughput benchmark performance levels are combined with the proposed availability level to determine a measure of overall proposed system-life throughput for the LSC and for the Analysis Computer. The actual throughput will be measured on a periodic basis, to be determined by the Government and Contractor, by combining the demonstrated benchmark performance with the operational use time on the LSC and on the Analysis Computer. The proposed performance levels must be met for each measurement of actual throughput regardless of past delivery of suites.

Shortfalls in throughput on the LSC or on the Analysis Computer are made up with new equipment brought in at no additional cost to the Government. Using the demonstrated benchmark performance on the upgraded LSC or Analysis Computer, the Government calculates how long the upgrade shall stay in place to compensate for the shortfall in throughput. This calculation is rounded up to a multiple of 6-month intervals to minimize disruption.

At the option of the Government, shortfalls in throughput on the LSC or on the Analysis Computer due to downtime shall cause downtime credits to accrue. These downtime credits shall be in lieu of bringing in new equipment. Downtime credits shall accrue on the HSMS or HFS.

To better reflect the Government’s computational needs over time, changes in the LSC and AC throughput benchmarks are made by mutual agreement between the Government and the Contractor throughout the life of the HPCS.

Accumulated computational cycles (in CPU-hours) that are lost when jobs are lost due to component failure or component reboot are not counted toward the system-life throughput calculation. If the accounting software cannot report the accumulated computational cycles, it will be assumed that 4 CPU-hours were lost for each processor on which a job ran.

All performance levels proposed for hardware and software upgrades must be met regardless of past delivery of suites

Uninterruptible power supplies (UPS) is required for all components of the HPCS. The UPSes provide sufficient power during environmental failures to gracefully shut down all components of the HPCS.

Support

The Government requires at least two software engineers (to provide a comprehensive system administration service), at least one hardware engineer, and at least one applications analyst available on site for at least eight hours within GFDL=s primetime window, five days per week. Additional on-call support is provided 24 hours per day, 7 days per week, with a 2-hour response time. The Government reserves the right to substitute hardware engineers with software engineers or application analysts during the life of the contract on an as-needed basis.

The Government requires an itemized list of all Contractor-supplied hardware and software items, and documentation of these items.

C.7.7 Benchmark Performance

C.7.7.1 A representative Climate IPCC Workstream

The following is an existing workstream for the IPCC. This is being provided to supply a better understanding of workstreams.

An example "modeling component" of the CM2.0 IPCC workstream is:

- The current target IT architecture for the workstream is the cluster of Origin 3000 computers located at GFDL, Princeton, NJ. See section C.7.1.3 for a more detailed description of the current target IT architecture for this workstream.
- Given the resource limits on the queues of the GFDL cluster, a typical simulation will produce 6 months of output in a little over 2 hours with a total throughput of about 5.2 simulation years per day.
- The model script startup utilizes 5 GB of input data (a combination of 0.5 GB of model restart files and 4.5 GB of model parameter files) moved from archive to local run directory.
- A model run produces 5 GB of restart files regardless of simulation length. The run script cpio's the restart files and copies them to archive.
- A 6 month simulation produces 5 GB of "raw" "history" model data spread across the 1350 files written by the 180 processes.
- Typically a script runs two 6 month simulations. At the end of each simulation, the raw history data is cpio'd, moved to archive. At the end of the script, a job is submitted to combine the raw individual process history data into global history data.
- The relationship between raw and global history data is fairly linear and so the 5 GB of raw, 6 month history produce about 5 GB of global history data covering 6 months.
- The global history output for a full year is then cpio'd and moved to archive.

- Thus, a simulation year produces about 10 GB of "raw" PE history data, 10 GB of global history data 1 GB of restart data.Raw and global history data and restarts are maintained in the archive.

The 10 GB of yearly global history data is the input to the post-processing component of the CM2.0 IPCC workstream.

An example "post processing" component:

- 100 years of global model history files (1 TB) are gathered to produce time series and averaged model output via post processing.
- During post processing, the 1 TB of global history is copied from archive to local disk in one hundred 10 GB files.
- Given the high cost of I/O and limitations of the netCDF file format, time series and averages for each of the atmosphere, ice, land and ocean model components are produced in separate file streams.
- For a 100-year run, 1 TB of global history input post processing produces 6.3 TB of output for further analysis.

C.7.7.2 The Representative Environmental Modeling Workstreams

C.7.7.2.1. The Environmental Modeling Test Bed activities, shown as WS4 above, represent weather forecast model research and development activities for the global, regional, and storm-scale forecast modeling activities. Typical activities include developing and testing new modeling and scientific technique, testing future upgrades to the current forecast modeling suite, and developing and testing new products and post processing techniques. The large quantities of model data used in this system originate from the real-time data assimilation system on the NOAA operational HPCS. Typical data usage and other characteristics are shown below.

C.7.7.2.2. Climate model development and calibration, shown as WS5 above is a computational intensive use of the computing resources. The coupled model system is run for about one simulated year every day for the past 25 years (or about 10000 total simulated years) in a hind cast mode to develop a detailed understanding of model performance. The extensive output is then analyzed to develop a well-founded calibration that is applied to the model when run in a forecast mode. Typical data usage generation and post-processing are shown below.

C.7.7.2.3 Data Assimilation development, shown as WS6 above, is both computational and data intensive use of the computing resources. Huge files of input data, along with model fields will be used to develop model-based analyses as a continuous function of short (1-6 hours) time windows. The huge input data streams originate from the real-time data assimilation system on the NOAA operational HPCS. Typical pre-processing, data usage, and post processing are shown below.

C.8 Appendix B – Available Government Facilities

This document provides descriptions of three government-furnished facilities that are designed to support high-performance computer systems and are available for possible use by Offerors to house systems under the NOAA HPC R&D contract, subject to the availability of facility resources as projected in Section C.8.11 below under the assumptions provided therein and subject to the facility terms and conditions provided in Section C.5.1**.** Two of these facilities, designated as BLDR-1 and BLDR-2, are located in the David Skaggs Research Center (DSRC), 325 Broadway, Boulder, CO, and the third, designated as PRTN, is located in Geophysical Fluid Dynamics Laboratory (GFDL) building complex, referred to as the “Princeton Complex” below, at 201 Forrestal Road, Princeton University Forrestal Campus, Princeton, NJ.

BLDR-1 currently houses NOAA’s JET computer system operated by OAR’s Forecast Systems Laboratory under a contract with HPTi. The BLDR-1 facility will not be available until October 2006. BLDR-2 is a facility currently under construction at DSRC and is expected to be available for use in October 2005. PRTN currently houses NOAA’s HPCS operated by OAR’s Geophysical Fluid Dynamics Laboratory under a contract with Raytheon.

C.8.1 Layout and Physical Dimensions of Computer Room

C.8.1.1 Computer Room Layout

BLDR-1
Figure 1 shows the computer room layout for BLDR-1. The area within the dotted line will be available in October 2006. This room is only designed for medium-density cooling configurations.

BLDR-2
Figure 2 shows the planned computer room layout, which is currently in the design phase. It will be designed with both overhead and under floor extreme-density cooling. The uninterruptible power supplies (UPS) will be fed from utility power only (i.e., the room will not be connected to the emergency motor generator). This room will be completed and available for use by the Offeror by October 2005.

BLDR-1 & BLDR-2 Building Diagrams
Offerors requesting building diagrams or related documents of federal government facilities shall be required to have an authorized representative execute a non-disclosure form as required under GSA Order PBS 3490.1. Specific room floor plans are available upon request and are not subject to these regulations.

PRTN
Figure 3 shows the computer room layout for the current system, but with the following planned modifications: move silos onto the northern hardpan and

construct five office cubicles on the southern hardpan. References to the front of the room in the following discussion refer to the bottom of the figure (nearest to the Operators Room), while the back of the room is at the top of the figure.

Figure 4 (attached) shows the overall Computer Building layout. The rooms shown at the bottom of the figure from left to right are:

- Loading Dock, which is designed to accept deliveries from 18-wheel trucks.
- Storage Room adjacent to the Loading Dock, which also serves as a receiving/staging area for deliveries to the Laboratory.
- Vendor Room, which currently provides office space for Contractor personnel of the existing system (This may be converted into office space with vendor offices relocated to the southern hardpan.)
- Operators Room, which serves as the control room for computer operations as well as security monitors for the Computer Room as well as outside building access points for the Computer Building and Main Building.
- Printer/User Output Room, which contains local Computer Building printers as well as user output bins.
- Operations Lounge.

C.8.1.2 Location of Computer Room and Characteristics of Surrounding Campus

BLDR-1 & BLDR-2

Both BLDR-1 and BLDR-2 are located within the David Skaggs Research Center (DSRC) in Boulder, CO. Due to the large number of visitors, the following website was developed to assist with locality information with regard to the site: http://boulder.noaa.gov/ This website will display maps, local information, driving directions, etc. Regarding specific locations: BLDR-1 is located in the "B" Block of the DSRC, on the 2nd floor, and BLDR-2 is located in the "A" Block of the DSRC, in the Garden level (basement).

PRTN

The Computer Room is in the Computer Building, which is one of two buildings in the PRTN Complex located at 201 Forrestal Road, Forrestal Campus, Princeton, NJ 08540 on Site B of Princeton University's Forrestal Campus. This campus is currently devoted to university research and is intended by Princeton University to be developed in the future as an office park on the U.S. Route 1 corridor. The nearest airport is the Trenton-Mercer Airport, which is roughly 12 miles away. The nearest highway, U.S. Route 1, is roughly 1/3 mile away.

C.8.1.3 Physical Dimensions of Computer Room

BLDR-1
Figure 1 indicates the computer room layout for BLDR-1. The entire computer room is 3600 square feet in size, with dimensions of 60 feet by 60 feet. The area within the dotted line in the figure will be available in October 2006 and is estimated to be approximately 2250 sq. ft.

BLDR-2
Figure 2 indicates the computer room dimensions and layout for BLDR-2. The entire computer room space is approximately 2100 sq. ft. A portion of this space is used for a substantial, Americans-with-Disabilities-Act (ADA)-compliant ramp for room access. 1500 sq. ft. of raised floor space is available for equipment. An adjacent room contains an additional 325 sq. ft. of raised floor space, separated from the main room by a wall.

PRTN
Figure 3 indicates the computer room layout for PRTN. Figure 4 shows the location of the computer room within the Computer Building of the PRTN Complex. The entire computer room is 10,004 square feet in size, with dimensions of 122 feet by 82 feet; this includes the UPS Room, which is located on the hardpan on the right rear corner of the Computer Room.

C.8.1.4 Raised Floor Space

BLDR-1
A total of 3600 sq. ft. of raised floor is in BLDR-1. 2250 sq. ft. of this space will be available in October 2006.

BLDR-2
A total of 1500 sq. ft. of raised floor will be in BLDR-2, all of which will be available once construction of the room is completed in October 2005.

PRTN
The raised floor area in PRTN totals 7052 sq. ft. with dimensions of 86 feet by 82 feet.

C.8.1.5 Non-Raised Floor Space and Equipment Staging Areas

BLDR-1 & BLDR-2
There is no non-raised floor in either computer room. The computer rooms themselves are used as staging areas.

PRTN
The non-raised floor on the left side of the room is 1476 sq. ft. with dimensions of 18 by 82 feet. The non-raised floor on the right side is 936 square feet, with dimensions of 18 by 52 feet, reflecting reduced non-raised floor space due to the presence of the UPS Room. Figure 3 indicates the planned future location of the five (5) StorageTek silos, four of which are currently located on the raised floor but will be relocated this winter to the locations on the hardpan as indicated in the figure.

C.8.1.6 Space for Vendor Personnel, Maintenance Space, and Vendor Storage

BLDR-1 & BLDR-2
The Government will supply two (2) 150-sq. ft. offices and approximately 100 square feet of lab space for use by the Contractor. None of this space will have furniture or lab benches. The space will have up to four (4) telephones as well as adequate LAN connections.

PRTN
Vendor support personnel (2 Computer Engineers and 2 Software Engineers) currently occupy three (3) rooms in the Computer Building. Equipment storage and maintenance space is provided on the left hardpan in the Computer Room. Plans are being made to construct five (5) offices on the left hardpan in the computer room to provide additional vendor office space while freeing up one of the current vendor rooms for other use.

C.8.1.7 *Height of ceiling above raised floor*

BLDR-1
Drop ceiling height = 8.5 feet

BLDR-2
Drop ceiling height = 10 feet

PRTN
Drop ceiling height = 9.5 feet

C.8.1.8 Maximum Allowable Height of Equipment

The following equipment height are mandated by fire codes and cooling limitations.

BLDR-1 & BLDR-2
Due to fire code and cooling limitations, the maximum allowable rack height will be 42U.

PRTN
Due to fire code regulations, the maximum allowable rack height will be no more than 18” below the drop ceiling height.

C.8.1.9 Space Below Raised Floor

The space below the raised floor is defined as height from the sub-floor base to the top of the raised floor.

BLDR-1
Raised floor height = 12 inches

BLDR-2
Raised floor height = 24 inches

PRTN
Raised floor height = 24 inches (except for the deeper chilled water trench, 4 feet deep, that extends down the middle of the room as indicated in Figure 4).

C.8.2 Installation Characteristics and Raised Floor Capacity

C.8.2.1 Physical Access for Equipment Installation

BLDR-1 & BLDR-2
Loading Dock - Standard, designed to accommodate 18-wheeled semi-trucks.

Access Path –

Corridors - 8” raised floors and 7-ft high doorframes

Elevator – Freight Elevator with 8000-lb. Capacity

BLDR-1: 25’ from Loading Dock to the freight elevator and then 50’ to the computer room

BLDR-2: 25’ from the Loading Dock to the freight elevator and then 150’ to the computer room, down and up ADA-compliant ramps.

PRTN
Loading Dock – Standard, designed to accommodate 18-wheeler semi-trucks.

Access Path – Equipment passes through an entry door from the Loading Dock into the Storage Room and then onto a hardpan staging area in the south corner of the Computer Room - a total distance of roughly 30 feet. The two sets of double doors along this path have clearances of 85 inches high by 70 inches wide.

C.8.2.2 Staging and Assembly Areas

BLDR-1 & BLDR-2
All staging and equipment assembly must be performed within the computer room.

PRTN
Equipment may be staged and assembled on the hardpan in the computer room immediately adjacent to the entry from the Storage Room and the Loading Dock. This non-raised floor area is at least 18’ wide by 20’ deep.

C.8.2.3 Loading Capacity of Raised Floor

BLDR-1
Computer Room Raised Floor Specs: ConCore SF 1250 Bolted Stringer: Concentrated Load: 1250 lbs.; Uniform Load: 300 lbs./ft²; Ultimate Load: 3850 lbs.; Rolling Load: 1000 lbs. (10 Passes)

Raised floor Specs (adjacent hallways): ConCore SF 1250 Cornerlock: Concentrated Load: 1250 lbs.; Uniform Load: 300 lbs./ft²; Ultimate Load: 3750 lbs.; Rolling Load: 1000 lbs. (10 Passes)

Floor tiles are 2’ x 2’ and are shown by location in the Figure 1 drawing. The dotted squares represent perforated tiles. Network and power connections are made via 5” round ports within the solid tiles only. All tiles, unencumbered by racks or other equipment, can be relocated for improved airflow, if necessary.

BLDR-2

Computer Room Raised Floor Specs: ConCore SF 1250 Bolted Stringer: Concentrated Load: 1250 lbs.; Uniform Load: 300 lbs./ft^2; Ultimate Load: 3850 lbs.; Rolling Load: 1000 lbs. (10 Passes)

Raised floor Specs (adjacent hallways): Concrete floors

PRTN

The raised floor in the Computer Room is designed to support a uniform live load of 250 pounds per square foot, with a deflection of not more than 0.040 inch. The raised floor is electrically interconnected to provide a common electric reference. Most of the raised floor, indicated as 2’-by-2’ square tiles in Figure 3, was replaced as part of the site preparation for the current Raytheon contract. The exceptions for this raised floor upgrade were for tiles under the following equipment: the two oldest StorageTek silos (when located on the raised floor), the MGE 250-kVA UPS, the Dataflow air-handler along the rear wall of the Computer Room, the Liebert 125-kVA PDU, and the EPE 125-kVA PDU.

Adjacent Non-Raised Floor and Pathway to Loading Dock: Concrete floors

C.8.3 Power Facilities

C.8.3.1 Power Service to Building

BLDR-1 & BLDR-2

480/277-volt main switchboard, rated for 3000 amps and 65,000 AIC. The gear has integrated IMPACC monitoring software. All power is fed from a single utility substation. Power is then distributed via step-down transformers and panel boards.

PRTN

A PSE&G utility substation with a 2500-kVA, 13200-to- 4160-volt, oil-filled transformer and 4160-volt switchgear provides electrical service to the Princeton Complex. This substation is located outside of the southwest corner of the Computer Building directly adjacent to the parking lot. This substation provides power support to the Princeton Complex as well as to several of the other buildings on the B site of the Forrestal Campus. Power usage for these other buildings is primarily for offices. PSE&G has indicated that it will upgrade the capacity of this substation, if required; however, the underground feeder line connecting the substation to the main PSE&G grid would be an additional cost to the customer.

An underground 4160-volt feeder, dedicated to the Princeton Complex, is routed from the utility substation to separate building substations located within the Princeton Main Building and Computer Building. The Main Building substation (1500-kVA 4160-to-480-volt transformer), located on the ground floor of Princeton's Main Building, provides power to the Main Building and to the Chilled Water Plant for the Princeton Complex.

The Computer Building substation is located in the Transformer Room, the location of which is shown in Figure 4. This substation is comprised of a 4160-volt air interrupter switch, a 1500-kVA silicone-filled transformer, and a 2000-ampere, 480/277-volt main switchboard. The equipment was installed around 1980 when the Computer Building was constructed and provides power to the Computer Building.

The lighting, large mechanical equipment (pumps, A/C units, etc), and some computer equipment are served at 480/277 volts via panel boards located throughout the Computer Building. The building receptacles, small motors, desktop computers, computer room equipment, and similar loads are served at 208/120 volts via step-down transformers and panel boards.

Currently all three chillers, described below, use the Main Building substation for power, a configuration that restricts their joint operation. Installation of an A-B switch for one or more of these chillers to permit use of power from the Computer Building substation, is a possible option for increasing available cooling capacity and providing increased flexibility for N+1 chiller backup.

C.8.3.2 Cost of Electrical Utilities Based on Recent Billing

BLDR-1 & BLDR-2

Electrical utility costs for DSRC occupants are currently charged to building tenants according to finished square footage occupied. Based on this algorithm, the electrical cost was $3.14 per square foot in FY 2004. There is a real possibility of metering electrical usage in the future. In this case, electrical costs would increase substantially. In FY 2004, the average cost for power was $0.05/kWH.

PRTN

The following table shows usage and expenditure data for the entire Princeton Complex as an indicator of recent electrical utility costs.

Total Usage and Expenditures for Electrical Utilities of the Princeton Complex for FY 1999 - FY 2004

	Annual Usage (kWH)	Expenditure	Avg. Cost / kWH
FY 1999	8,826,400	$713,948	$0.0810
FY 2000	9,032,000	$640,544	$0.0707
FY 2001	8,555,200	$644,035	$0.0756
FY 2002	5,976,000	$431,506	$0.0722
FY 2003	6,593,600	$472,405	$0.0711
FY 2004	7,981,571	$654,619	$0.0820

C.8.3.3 Power conditioning and UPS capabilities

BLDR-1

- Power Distribution: Cutler-Hammer Electrical Distribution Equipment (480Volt, 3 Phase) Under floor distribution is accomplished by 50 ft. flexible conduit power whips, fed from wall-mounted breaker panels.
- UPS: 300 kVA Chloride UPS Systems (450 kVA Installed); (250kVA available to R&D HPCS in 2006); 8-Minute Runtime (Full Load)
- Other: Transient Voltage Surge Suppressor (TVSS) Protected
- Emergency Power Off (EPO) Switch Protected

BLDR-2

- Power Distribution: Cutler-Hammer Electrical Distribution Equipment (480Volt, 3 Phase)
- Under floor power distribution will be accomplished by fixed receptacles attached to rigid conduit and mounted one foot below the floor tiles.
- UPS: 350 kVA UPS System (500 kVA Installed); 16-Minute Runtime (Full Load)
- Other: Transient Voltage Surge Suppressor (TVSS) Protected
- Emergency Power Off (EPO) Switch Protected

PRTN

- Power Distribution Units (PDUs) [shown in dark blue in Fig. 3]:
- Three (3) United Power 225-KVA PDUs (owned by Raytheon)
- Liebert 125-KVA PDU (owned by Government/Princeton University)
- EPE 125-KVA PDU (owned by Government/Princeton University)
- UPS [shown in brown in Figure 3]:
- 500-kVA UPS MGE cabinet and battery bank (owned by Raytheon)
- 225-kVA UPS MGE cabinet and battery bank (owned by Government/Princeton University)

C.8.3.4 Backup Power Generator Capabilities

BLDR-1

- All Equipment fully backed up by emergency backup generators.

- Current Generator is 1250-KW Cummins Diesel. 1800-gallon tank capacity allows 24 hours of operation

BLDR-2
The DSRC EM Generator has reached its rated capacity, and no further load will be allowed. Therefore, the BLDR-2 facility will not have EM generator backup. The UPS systems will be designed to sustain a 16-minute outage at full load. The recent power loss history at DSRC is presented in Section 8a below.

PRTN
Natural-gas-fired, 75-kW generator: activates automatically when power is lost. This generator is only designed to allow pumps within the chilled-water cooling system to continue to run during a brief power outage.

C.8.4 Cooling Facilities

C.8.4.1 Cooling Service to Building

BLDR-1 & BLDR-2
Building operates three (3) 470-ton chilled water cooling systems. Two of the systems are on EM generator power, and one is fed strictly from utility. There are three cooling towers, each sized to match the capacity of a single chiller. There are three primary pumps, three secondary pumps, and three condenser water pumps. The chiller plant is located in a mechanical room that is adjacent to BLDR-2. Cooling is delivered at 42^{o}F. There is a rough-in plan for an additional chiller and space for an additional cooling tower in the future, but no plans have been authorized to expand the chiller plant at this time. The chiller plant is currently operating at full capacity.

PRTN
Primary Chilled Water Plant, located in the mechanical room and tower bay southeast of the Transformer Room (see Figure 4), contains two chillers:
- 400-ton Carrier centrifugal chiller (designated Chiller #1)
- 350-ton York centrifugal chiller (designated Chiller #3)
- Baltimore Air Coil Cooling Towers

Chiller #1 was installed in the spring of 2000 along with new cooling towers and pumps. Chiller #3 was installed in 1996.

Secondary Chilled Water Plant, located in the Transformer Room (see Figure 4):
- 225-ton Carrier centrifugal chiller (designated Chiller #4)
- Baltimore Air Coil Cooling Tower

Chiller #4 was installed in 1979. Cooling tower replacement will be completed by November 2004.

The Primary Chilled Water Plant provides cooling to the entire Princeton Complex. The Secondary Chilled Water Plant is designed to only support the Computer Building.

The current configuration only permits one chiller in the Primary Chilled Water Plant to operate at one time. These restrictions are due to limitations on power usage (i.e., all chillers run off of the Main Building substation – see above) and to flow restrictions caused by the current chilled piping design. Chiller #4 has only been used on especially hot days for emergencies, or as backup to partially support the cooling load during periods of time when one of the primary chillers is taken off-line for servicing or repair. However, the cooling tower for the Secondary Chilled Water Plant has been replaced during the fall of 2004. Because of this enhancement, Chiller #4 is expected to be capable of providing a more robust cooling load.

Cooling is delivered to the Computer Room through a six-inch and eight-inch piping system from the mechanical room at a temperature of 42^{o} F, plus or minus 2 degrees. The pipe enters the computer room in a trench that is 4 feet deep under the raised floor in the center of the computer room, as indicated in Figure 4. The piping to the Computer Building currently limits the flow of chilled water to the Computer Room and will require redesign if additional cooling capacity is required.

C.8.4.2 Room Air Conditioning Capabilities

BLDR-1

- 90-Ton Liebert Downdraft Computer Room Air Conditioners (CRACs) (De-rated for altitude) (120 Tons Installed via 4 x 30 Ton units)
- Cooling powered from Emergency circuits.

BLDR-2

- 90-Ton Liebert Downdraft CRAC (De-rated for altitude) (120 Tons Installed via 4 x 30 Ton units)
- Thirteen (13) 4-Ton Liebert Extreme Density Overhead (XDO) Systems
- Cooling powered from UPS

PRTN

- Seven (7) 35-ton Dataflow/APC CRACs (owned by the Government/ Princeton University)
- Other smaller air handling units - located in the ceilings of the Printer Room and the UPS room

The primary source of humidification for the Computer Room is an air conditioner and steam boiler, located in the building tower on the western corner of the Computer Building.

C.8.4.3 Availability of plumbing for chilled water

BLDR-1 & BLDR-2

All chilled water plumbing is installed to accommodate the Computer Room Air Conditioner (CRAC) units and the Extreme Density Overhead systems. No additional plumbing capability or capacity is available.

PRTN
The computer room has operated several water-cooled computer systems (most recently - Cray T932, T94, and T3E) and thus has been previously configured to support water-cooling, including a chilled-water trench running down the middle of the raised-floor area of the computer room, as indicated in Figure 4.

C.8.5 Networking Facilities

C.8.5.1 Wide-Area Networking Services To Computer Room

BLDR-1 & BLDR-2
NOAA Boulder maintains multiple wide-area networking links. These include a one-gigabit per second (1 Gb/sec) connection shared with NCAR from Boulder to the Denver Front Range GigaPoP (FRGP) via dark fiber. This link provides connections to the commercial Internet via AT&T, Cable & Wireless, and Level3 at 20 Mb/sec for each path. The NOAA/NCAR-FRGP link also provides a 310-Mb/sec connection to Internet2. In addition, NOAA Boulder maintains a secondary 20-Mb/sec connection to the commercial Internet through the University of Colorado (CU). CU also maintains a separate 622-Mb/sec link to the FRGP that is available to NOAA Boulder should the NCAR link fail.

PRTN
PRTN maintains a 9 Mb/sec connection to the commercial Internet under its current Raytheon contract. In addition, the B-Site Forrestal Campus maintains a 100-Mb/sec microwave connection to the Main Campus of Princeton University. This connection provides access to Princeton University's 100 Mb/sec connection to Internet2.

C.8.5.2 Proximity to additional wide-area networking capability

BLDR-1 & BLDR-2
NOAA Boulder is a joint owner of the Boulder Research and Administrative Network (BRAN). BRAN is an eleven-mile dark fiber metropolitan-area network owned by NOAA, UCAR/NCAR, CU, and the City and County of Boulder. Besides interconnecting BRAN owner sites, it also has key tactical access to commercial communication providers, US West and ICG. NOAA has indisputable access to 24 strands of BRAN fiber. NOAA Boulder and the FRGP have contracted with the National Lambda Rail (NLR) network to attach in fiscal year 2005. Using Dense Wavelength Division Multiplexing, NLR will provide multiple channels of service at startup including: a 10-Gigabit Ethernet routed circuit to Internet and Internet2, multiple one-Gigabit Ethernet switched circuits to NLR nodes, multiple wavelengths for dedicated point-to-point circuits, and service for the Global Lambda Integration Facility.

PRTN
Level3 maintains a dark-fiber run along U.S. Route 1, which is roughly 1/3 mile from the lab across the field. The B-Site of Forrestal Campus is connected to the DOE-supported Princeton Plasma Physics Laboratory (PPPL) through dark fiber.

PPPL is on DOE's ESnet wide-area network with a bandwidth capacity of at least OC3.

C.8.6 Fire Alarm And Fire Suppression Capabilities

C.8.6.1 Fire Alarm and Suppression Systems

BLDR-1 & BLDR-2
Both computer rooms will have FM-200 fire suppression with VESDA fire detection systems. Both will have wet-pipe overhead sprinklers with 155^{o}F trip point. CO_2 bottles are located in the room with FE-36 bottles surrounding the room. Alarms are tied to the building alarm system.

PRTN
The PRTN computer room currently is protected by a high-voltage fire alarm system and a wet-pipe water sprinkler system. However, this winter, these systems will be replaced with the following: (1) a low-voltage alarm system with advanced alarm panel in the room adjacent to the computer room, (2) a dry-pipe water sprinkler system, (3) plumbing and other infrastructure to support future use of a gaseous fire suppression system (such as FM-200), and (4) a VESDA early smoke detection system.

C.8.6.2 Fire Extinguishers

BLDR-1 & BLDR-2
Fire extinguishers located within the computer rooms are CO_2 bottles. Fire extinguishers located within the hallways and computer room access areas are FE-36 gas bottles.

PRTN
Following the NCEP fire event in 2000, PRTN reviewed its fire procedures and eliminated any dry-chemical fire extinguishers from the entry paths to the Computer Room. The lab also met with the supporting fire departments, both Plainsboro and Princeton Plasma Physics Laboratory (PPPL) fire units, to explain the dangers of dry chemical extinguishers to computer equipment and to verify fire suppression procedures to be followed during fire emergencies.

C.8.6.3 Response Time Of Local Fire Department

Response time is defined as the time duration between fire alarm activation and the arrival of fire department personnel and equipment at the site.

BLDR-1 & BLDR-2
Boulder Fire Department provides fire-fighting capability. A station is located approximately one mile away and the response time is between 4 to 6 minutes.

PRTN
The Princeton Complex has measured the response time of the fire department (PPPL as first responders, backed up by Plainsboro Fire Department) to average

10 minutes (minimum of 7 minutes and maximum of 14 minutes), based on 6 events (2002-2004). These events were caused by unrehearsed false alarm events that were caused by faulty alarm sensors. [The Princeton Complex will be replacing the entire fire alarm system during the winter of 2004-05.]

C.8.7 Special Capabilities

C.8.7.1 Capabilities for Continuous Operation

The following describes any capability that the facility may have to permit continued operation during a power outage or to ride through power anomalies.

BLDR-1
All electrical systems are fully backed up for uninterrupted service to cover for any length of power outage. The CRAC units are powered from emergency circuits, and will experience a shutdown until the EM generator comes online and transfers power (usually one to two minutes.) at which time they will automatically and sequentially restart. Once the EM generator starts, it will continue to operate until clean and stable utility power is maintained for 30 minutes, at which time the generator will transfer the load back to utility power and shut down.

BLDR-2
No EM generator capacity is available for this facility; therefore, all equipment will be placed on to UPS power fed from utility. Should a lengthy outage occur, the UPS would provide uninterrupted service, until which time a graceful shutdown of computer equipment is required. CRAC units and cooling systems will remain on until all equipment is off and all static air has been forced from the compute nodes. The CRAC units will then shutdown, prior to exhausting the battery capability of the UPS systems.

The DSRC outage history is in the following table and shows one outage per year that would mandate a shutdown of the computer room.

Recent History of Power Loss Episodes at DSRC

FY2003	
DATE	DURATION
26 Mar 03	13 Minutes
13 May 03	13 Minutes
8 Sep 03	30 Seconds
11 Sep 03	2 Minutes
FY2004	
10 Jun 04	52 Minutes

PRTN
PRTN's facility environment is designed primarily for riding through a brief power outage, assuming the outage is sufficiently brief that the chillers will be able to automatically restart so that the temperature and humidity conditions in the computer room are maintained within acceptable limits.

The following table indicates environment related outages for the last three fiscal years. During these years, there were 29 documented power fluctuation episodes (14 in FY02, 11 in FY03, and 4 in FY04) in which the UPSs and PDUs maintained conditioned power to the computer systems and the chillers either continued to operate or else shutdown and then restarted automatically.

Recent History of Environment-Related Outages for PRTN

FY2002	
23 Jul 02	Mechanical (chiller) outage of 53-minute duration with 36 minutes of system recovery time
FY2003	
25 Feb 04	Power outage of 103-minute duration with 224 minutes of system recovery time
FY2004	
	No outages

C.8.7.2 Duration Of Ride-Through At Full Load

BLDR-1
Not Applicable – Maintain operation during power outage

BLDR-2
Designed for graceful shutdown during a power outage lasting longer than 10 minutes, with CRAC units shutting down last.

PRTN
The UPS systems for the current system provide a ride-through for the computer systems of 30 minutes on the 500-KVA UPS and 60 minutes on the 225-KVA UPS. However, the room air conditioning units and the chillers do not have generator support. If the power interruption is brief, Chillers #1 and 3 are likely to restart automatically. This recycling capability, combined with generator backup for the chilled water pumps, permits the facility to ride through brief power interruptions without the system going down.

C.8.7.3 Capabilities for Remote Facility Management

The following describes any capability that allows the facility to be monitored and managed remotely.

BLDR-1 & BLDR-2
BLDR-1 has SCADA (Supervisory Control and Data Acquisition) coverage for environmental monitoring. BLDR-2 is projected to have SCADA coverage as well. CRAC units also have building automation tie-in for alarms. Both systems have dial-out capability.

PRTN

Currently the computer room does not have remote monitoring capabilities. Such capabilities are being investigated for possible future implementation.

C.8.7.4 Capabilities for Lights-Out Operation

The following describes any capability that allows the facility to operate during off-hours without operators present.

BLDR-1 & BLDR-2
BLDR-1 currently has no operator coverage from 7pm to 7am every day. The SCADA environmental monitoring system will trigger an automated Emergency Power Off (EPO) if the temperature reaches 94oF.

PRTN
Currently the Computer Room does not have remote monitoring capabilities. Such capabilities are being investigated for possible future implementation.

C.8.8 Physical security

C.8.8.1 Campus Security

BLDR-1 & BLDR-2
Both facilities are located on federal property with federal police providing access to the grounds round the clock.

PRTN
Security personnel from the Princeton Plasma Physics Laboratory patrol the B-Site of the Forrestal Campus on a regular basis during off-hours. Plainsboro police are also available on call. Key sections of the exterior buildings are lighted at night. Security cameras observe these areas, digitally recording and displaying on monitors in the Operations Room on a round-the-clock basis.

C.8.8.2 Security of Building Containing Computer Room

BLDR-1 & BLDR-2
Federal police stationed at main entrance only during work hours. All entrances are monitored by video surveillance at all times.

PRTN
DOC Building Access Policies, involving NOAA ID cards and temporary badges, prescribe access controls for the Princeton Complex. NOAA issues these IDs according to prescribed procedures. During off-hours, access is limited to one building entrance. Building entrances are monitored from the Operations Room by a camera monitoring system with a digital recording system. The Operations staff carries out building security tours on a frequent schedule to monitor building environment, security, and facilities infrastructure.

C.8.8.3 Computer Room Security

BLDR-1 & BLDR-2

Computer room has restricted access provided by cipher locks at all entrances. A proximity badge and/or PIN number are required at all times.

PRTN
Computer Room access is controlled by a centrally managed cipher lock system to restrict room access to approved personnel. Sign-in procedures are required for individuals not on the approved personnel list. Cameras in the Computer Room allow Operators to monitor and digitally record areas of the room not viewable from the window of the Operations Room.

C.8.8.4 Procedures For Contract Personnel To Access A Government Facility

BLDR-1 & BLDR-2 & PRTN
All contract personnel must submit to a background check and will be issued an "affiliate" NOAA ID badge.

C.8.9 Estimated Annual Facility Cost of Operation of Computer Room (Excluding Power Costs)

BLDR-1 & BLDR-2
Annual costs for rent, maintenance, heating, and other general services were $42.25/sq. ft. in FY 2004 as paid to the building landlord, GSA.

PRTN
Annual facility costs, excluding electrical utility costs, are estimated to be $10/sq. ft. This includes: rent paid to Princeton under the Government triple-net lease; facilities services (physical plant maintenance, equipment service, janitorial services, etc.); utilities (other than electrical power); and wide-area network fees.

C.8.10 Availability of blueprints

BLDR-1 & BLDR-2
Blueprints are available from local GSA office and will need to be signed for by approved contractors with a need to review them.

PRTN
The Princeton Complex maintains a hardcopy archive of all blueprints for the Computer Building and associated chilled water plant upgrades.

C.8.11 Projected Availability of Floor Space, Power, Cooling, and WAN Bandwidth

The following amounts show projected total resources available at the indicated dates under the assumptions provided. The October 2004 availability is shown for reference purposes only, since Contractors cannot utilize the resources until October 2005.

		Oct. 2004	Oct. 2005	Oct. 2006
Available Raised Floor Space (sq. ft.)				
	BLDR	0	1,500	3,750
	PRTN	1,900	2,700	5,500
	TOTAL	***1,900***	***4,200***	***9,250***
Available Non-Raised Floor Space (sq.ft.)				
	BLDR	0	0	0
	PRTN	360	360	360
	TOTAL	***360***	***360***	***360***
Available Power for New Hardware (kVA)				
	BLDR	0	350	600
	PRTN	250	150	575
	TOTAL	***250***	***520***	***1,175***
Available Cooling for New Hardware (tons)				
	BLDR	0	146	218
	PRTN	180	155	283
	TOTAL	***180***	***301***	***501***

Assumptions for Projections

Available Raised Floor Space:

BLDR:
(1) October 2005 estimate assumes completion of computer room design, construction and testing.
(2) October 2006 estimate reflects available space occupied by systems purchased under current HPTi contract. In reality, the new contractor will be required to coordinate new equipment installation with current system removal so as to minimize loss of compute cycles to the Government.

PRTN:
(1) Oct. 2004 estimate assumes StorageTek silos have been moved onto non-raised floor as indicated in Fig. 3.
(2) Oct. 2005 estimate assumes the mid-contract upgrade under the current Raytheon contract will provide 20% of additional free raised floor space compared to current floor-space usage.
(3) Oct. 2006 estimate reflects available space occupied by systems leased under current Raytheon contract. In reality, the new contractor will be required to coordinate new equipment installation with current system removal so as to minimize loss of compute cycles to the Government.

Available Power:

BLDR:
(1.) The projected available power is dictated by the cooling capacity that is available. Currently, the chilled water plant has reached its rated capacity and no further load can be supported without the addition of a new chiller.

PRTN:
(1.) The projected available power assumes, somewhat arbitrarily, that total power to the computer room is limited to the current UPS capacities, which total 725 kVA, reduced by a 10% safety factor to roughly 650 kVA. This is a conservative estimate, given the total substation capacity, but does not include usage of chillers, which are currently run off the Main Building substation.
(2.) Power estimates assume steady state ratings, not book value ratings for systems.
(3.) Oct. 2005 estimate assumes a 25% increase in power usage over current usage.
(4.) As with floor space, the Oct. 2006 reflects only minimal system usage from the current system with a background usage of 75 kVA.

Available Cooling:

BLDR:

(1.) The chill water plant has reached its maximum capacity and will lose its N+1 capability on hot days. A failure within the chill water plant, on a hot day, will force the BLDR-2 computer room to load shed its cooling resources and shut down.

PRTN:

(1.) Assume the building load for the Princeton Complex for the hottest day of the year is 274 tons, based on observation that Chiller #1 alone (400 tons) will reach maximum capacity on this day when cooling the Princeton Complex plus the current system.

(2.) Compute cooling load for the system by multiplying system power usage (converted to tons) by 1.3 to reflect (conservatively) other room heat sources and cooling inefficiencies. (3.) Assume maximum chiller capacity uses chillers #3 (350 tons) and #4 (230 tons) operating together. Joint operation of chillers #1 (400 tons) and #4 (230 tons) are also viable within the current configuration. However, as indicated above, joint operation of chillers #1 and #3 will require electrical and mechanical (plumbing) modifications.

		Oct. 2004	Oct. 2005	Oct. 2006
Available Wide-Area Network Capacity (Mb/sec.) (list by type below)				
Abilene / Internet 2	BLDR	310	500	0
Commercial Internet		80	80	100
National Lambda Rail (NLR)				
NOAA Research / Internet 2		0	1,000	10,000
Commercial Internet		0	1,000	1,000
***Note: 10,000 Mbps and/or dedicated Lambda's (wavelengths) available via NLR in 2005 and 2006 if needed.**				
	PRTN			
Microwave to Internet2 (PU allowed)= 50 Mb/s		25	20	15
Commercial Internet = 9 Mb/s				
	TOTAL	***415***	***2,600***	***11,115***

Assumptions for Projections

Available Wide-Area Network Capacity:

BLDR:

(1.) WAN capacity assumes timely completion of National Lambda Rail installation and connectivity

PRTN:

(1.) Assume Princeton University only permits NOAA to use a total of 50% of its Internet2 bandwidth capacity.
(2.) Assume current NOAA usage of total bandwidth will grow at a rate of 15% per year.

[Figure 1 withheld from Public Website]

Figure 1 *Projected Computer Room Layout for BLDR-1 in September 2006. The area enclosed within the dashed line indicates the floor space that will be available for Offeror use in October 2006. The small squares shown indicate 2'x2' floor tiles. Vent tiles are indicated by stippled squares.*

[Figure 2 withheld from Public Website]

Figure 2 *Schematic of Computer Room Layout for BLDR-2 facility that is expected to be available in October 2005*

[Figure 3 withheld from Public Website]

Figure 3 *Computer Room Layout for PRTN as of January 2005. Small squares in the figure indicate 2'x2' floor tiles. Vent tiles are indicated by stippled squares.*

[Figure 4 withheld from Public Website]

Figure 4 *Computer Building Schematic for Princeton Complex*

C.9 Appendix C – Terms

C.9.1 Abbreviations

ADIC	Advanced Digital Information Corporation
ANSI	American National Standards Institute
CM	Configuration Management
CM2	GFDL's Climate Model Version 2
CONOPS	DOC's Concept of Operations acquisition process
COTS	Commercial Off-The-Shelf
CRAC	Computer Room Air Conditioner
CXFS	SGI's shared filesystem for storage area networks (SANs)
DDN	Direct Data Networks
dGB	Decimal Gigabyte (10^9 bytes)
dPB	Decimal Petabyte (10^15 bytes)
DOC	US Department of Commerce http://www.commerce.gov
DSRC	David Skaggs Research Center
ENTA	Enterprise Network Target Architecture
FC	Fiber Channel
FISMA	Federal Information Security Management Act of 2002
FSL	NOAA's Forecast Systems Laboratory http://www.fsl.noaa.gov 325 Broadway Boulder, CO 80305
FYxxxx	Fiscal Year xxxx
Gb	Gigabit (2^30 bits)
Gb/s	Gigabit(s) per second
GB	Gigabyte (2^30 bytes)
GFDL	NOAA's Geophysical Fluid Dynamics Laboratory http://www.gfdl.noaa.gov Princeton University Forrestal Campus 201 Forrestal Road Princeton, NJ 08542
GFE	Government Furnished Equipment
GNU	Unix-compatible software system
GPFS	General Parallel File System
GrADS	Grid Analysis and Display System
GSA	General Services Administration
HBA	Host Bus Adapter
HFS	Home File System
HIMF	Hallberg Isopycnal Model (Fortran)
HPC	High Performance Computing
HPCS	High Performance Computing System (see definition)
HPSS	High Performance Storage System

HSI	Hierarchical Storage Interface
HSM	Hierarchical Storage Management
HSMS	Hierarchical Storage Management System
htar	HPSS tape archiver
IDL	Interactive Data Language
IEEE	Institute of Electrical and Electronics Engineers
IMSL	International Mathematical and Statistical Library
I/O	Input/Output
IPCC	Intergovernmental Panel on Climate Change
IT	Information Technology
ITS	Information Technology Services
KB	Kilobyte (2^10 bytes)
kVA	Kilovolt amperes (1000-volt amps).
LAN	Local Area Network
LPAR	Logical Partitioning
LSC	Large-Scale Computing
LTD	Live Test Demonstration
Mb	Megabit (2^20 bits)
Mb/s	Megabit(s) per second
MB	Megabyte (2^20 bytes)
MPI-2	Message Passing Interface (latest version)
MTBF	Mean-Time-Between-Failure
NAG	Numerical Algorithms Group
NCAR	National Center for Atmospheric Research
NCEP	NOAA's National Centers for Environmental Prediction http://www.ncep.noaa.gov 5200 Auth Road Camp Spring, MD 20746
NetCDF	Network Common Data Format
NIST	National Institute of Standards and Technology
NLR	National Lambda Rail
NOAA	DOC's National Oceanic and Atmospheric Administration http://www.noaa.gov
NWS	National Weather Service
OAR	Office of Oceanic and Atmospheric Research
OCIO	Office of the Chief Information Officer
OMB	Office of Management and Budget
OPM	Office of Personnel Management
PB	Petabyte (2^50 bytes)
PDU	power distribution units
PFTP	Program to transfer data from host to host
PM	Preventive Maintenance
POSIX	Portable Operating System Interface
PSE&G	Public Service Gas and Electric
R&D	Research and Development

RAID	Redundant Arrays of Inexpensive Disks
RFI	Request For Information
RFP	Request For Proposal
RUC	Rapid Update Cycle
SAN	Storage Area Network
SGE	Sun Grid Engine
STK	Storage Technology
TB	Terabyte (2^40 bytes)
UDUNITS	Unidata's Units Library
UPS	Uninterruptible power supplies
WAN	Wide Area Network
WRF	Weather and Research Forecast
WRF-CHEM	Weather and Research Forecast with Atmospheric Chemistry
WRF-EM	Weather and Research Forecast – Eulerian Mass
WRF-NMM	Weather and Research Forecast – NCEP Mesoscale Model

C.9.2 Definitions

Application Memory	The maximum resident set size used by an application process in Mega/GigaBytes.
Availability	The availability level of a computer, component, or device is a percentage figure determined by dividing the operational use time by the difference between wallclock and null time.
Communication Fabric	The hardware component(s) supporting MPI message traffic.
Community Supported Software	Open Source Software
Degraded Mode	System operation at less than normal capability due to the loss of hardware or software components on that system.
Downtime	That period of time when all of an HPCS component's workload cannot be accomplished due to a malfunction in the Contractor-maintained HPCS hardware or software, or when the HPCS or a component of the HPCS is released to the Contractor for maintenance.
HPCS	A High Performance Computing System can be one large System or an aggregation of sub-Systems. A given System/sub-System can be further described by individual components. A single component may span multiple sub-Systems.
Null Time	The period of time when the workload cannot be accomplished due to environmental failure, such as loss of electric power or cooling, or recovery from environmental failure.

Operational Use Time	The time during which equipment is available to the Government, exclusive of preventative maintenance time, remedial maintenance time, or Contractor-caused machine failure. Partial credit may be given by the Government for equipment operating in degraded mode (for example, when a portion of the processors, memory, disk, etc. on a computer is unavailable). The Government may declare the entire HPCS down even if parts of the HPCS are available.
Physical Memory per Processor Core	The size of the memory chipset in Mega/GigaBytes supporting a single processor core.
Preventative Maintenance (PM)	That maintenance performed by the Contractor which is designed to keep the equipment in proper operating condition. It is performed on a scheduled basis.
Processor Core	The component of the processor containing the dependently scheduled floating point and integer registers and arithmetic and load/store unit(s).
Processor Socket	The motherboard component designed to receive the processor.
Remedial Maintenance (RM)	That maintenance performed by the Contractor which results from Contractor-supplied equipment or operating software failure. It is performed as required and therefore on an unscheduled basis.
Suite	A set of concurrent instances of a workstream and possibly, a number of given workstreams (see target IT architecture).
Target IT Architecture	The set of hardware and software systems which accomplish a workstream suite.
Total (Workstream) Throughput Time	The wallclock time from submission of the first instance of a workstream component to the successful program end of the last instance of a workstream component. In general, multiple workstream types and instances may be running on a particular target IT architecture. The Total Throughput Time for a workstream is defined in the context of all concurrent loads for the target IT architecture.
Workstream	A single instance of end-to-end processing.